AFGL-TR-79-0058

LEVEL II

12

IMPROVED DETERMINATIONS OF THE EARTH'S GRAVITY FIELD

Georges Blaha

Nova University Ocean Sciences Center
8000 North Ocean Drive
Dania, Florida 33004

Scientific Report No. 1

31 January 1979

Approved for public release; distribution unlimited

AIR FORCE GEOPHYSICS LABORATORY
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
HANSCOM AFB, MASSACHUSETTS 01731

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: AFGL-TR-79-0058	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): IMPROVED DETERMINATIONS OF THE EARTH'S GRAVITY FIELD,	5. TYPE OF REPORT & PERIOD COVERED: Scientific ~~Report No.~~ 1
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Georges Blaha	8. CONTRACT OR GRANT NUMBER(s): F19628-78-C-0013
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Nova University Ocean Sciences Center, 8000 North Ocean Drive, Dania, Florida 33004	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F, 2309G1AP, G1
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Geophysics Laboratory, Hanscom AFB, Massachusetts 01731, Contract Monitor: George Hadgigeorge/LWG	12. REPORT DATE: 31 January 1979
	13. NUMBER OF PAGES: 160
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 162p.	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Satellite altimetry	Least squares adjustment
Gravity anomaly	Constraints
Geoid undulation	Potential
Spherical harmonics	Point mass

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The main thrust in the present report is directed toward improving the knowledge of the earth's gravity field and of its fundamental surface, the geoid. The vehicle for research along these lines is the concept developed in recent years, based on a combined adjustment of the short arc model of satellite altimetry and the gravity anomaly model. The improvements are realized in two basic categories, by upgrading the accuracy of the algorithms used previously, and by providing for a more detailed representation of the local

(Continued)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

393 866 LB

geoid in areas with a dense configuration of altimeter observations.

The first two improvements belonging to the first category are concerned with adjustments of satellite altimeter data. A quantity needed in this process is the radial distance (length of a position vector) from the geocenter to the initial geoid. In the past, this distance has been computed to a desired accuracy in an iterative process. Even if the best possible starting value is taken, such as the radial distance to the mean earth ellipsoid, two iterations are still needed. However, an algorithm has been designed which yields a subcentimeter accuracy already from the first iteration. In most practical cases where an iterative solution would not be attempted for economical reasons, this new algorithm symbolizes a significant increase in accuracy.

A key element in any attempt to achieve a high accuracy (perhaps to a decimeter) in geoid representation via satellite altimetry is the necessity either to obtain the ephemeris of comparable accuracy, or to circumvent this requirement by adjusting the ephemeris in some way, together with the geoid. The second approach, adopted in this study, allows for a piece-wise treatment of short orbital arcs considered mutually independent. The main question which had to be answered when pondering the possibility of using the short arc adjustment model in altimetry reductions is whether or not this method is inherently capable of representing the geoid to the desired accuracy. The contained analysis provides at least a partial answer to this question by pointing out the necessary conditions for accomplishing this goal. They consist almost exclusively in specifying the maximum length of a satellite arc compatible with the given accuracy requirements.

Another improvement in accuracy is realized when the foundation of the gravity anomaly model is analyzed. A basic equivalence is noticed, at the level of observation equations expanded in terms of spherical harmonic coefficients, between the gravity anomaly model and the gravity model. However, a close scrutiny reveals the following fine distinction. The gravity anomaly model contains two approximations, called the spherical and the direction approximations. The gravity model, on the other hand, contains only the direction approximation which is the less important of the two. The accuracy of the gravity anomaly model is thus increased upon transforming it to the corresponding gravity model.

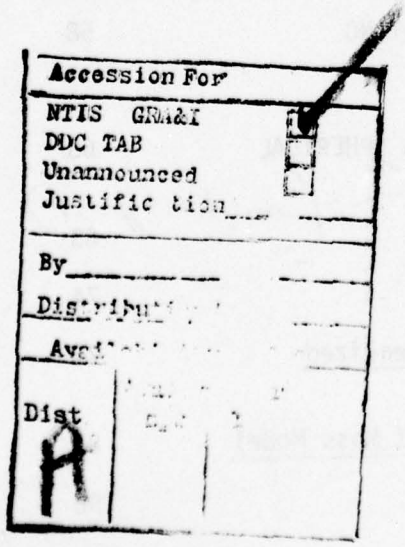
The purpose of introducing the concept of point mass parameters when treating gravity anomalies and satellite altimetry is to add fine local structure to a geopotential model based on the spherical harmonic coefficients. The point masses are to be employed only in areas of a relatively high observational density, thus assuring that the number of parameters may be kept reasonably small. Originally, two kinds of constraints, called mass constraints and volume constraints, had been assumed to participate in the point mass adjustment. However, in the practical cases of modeling geoid undulations (derived from an adjustment of satellite altimetry) in one area, no constraints appear necessary. On the other hand, one mass constraint, which assures that the sum of all the point mass magnitudes introduced in this adjustment is zero, may be beneficial for the predictions of gravity anomalies. Further results and suggestions regarding the density of the point masses, their depth, etc., are presented.

Partial

TABLE OF CONTENTS

(Cont on p. iv)

TABLE OF CONTENTS (CONTINUED)



1. INTRODUCTION

Over the last few years, AFGL has developed methods and software, documented in several reports and papers, whose purpose is the determination of the earth's gravity field. The determination of the spherical harmonic potential coefficients is an example related to the global gravity field and to the shape of the geoid on a worldwide basis. The observables (i.e., the quantities that can be observed) in the mathematical models used for such purpose are typically satellite altimeter data gathered over the oceanic regions, and mean gravity anomalies available for most of the continental regions. Clearly, sea gravity anomalies or direct gravity values could be used equally well. The satellite altimeter measurements result in the distances between a satellite at given times and the geoid.

Traditionally, approaches to the reduction of satellite altimetry have contained a requirement that extremely accurate reference orbits be established from satellite tracking. For example, in order to satisfactorily represent the geoidal surface in the "long arc mode" of satellite altimetry, the radial component of each satellite position should be determined perhaps tenfold better than what can be reasonably expected from the existing global tracking networks. In this mode, the trajectory of an orbiting satellite is considered in its entirety as a continuous, uninterrupted curve whose size, shape, orientation, etc., are dictated by the dynamical considerations.

The approach undertaken by AFGL in the earth's gravity field determination using satellite altimetry takes advantage of enormous run-time and core space computer savings associated with the "short arc mode." An equally important advantage embodied in this mode is that highly precise reference orbits are no longer a stringent requirement. The reference orbits, initially established from some routine global tracking, e.g., the VHF Doppler tracking, are divided into a large number of sub-arcs (they may number in the thousands in one adjustment process). Each sub-arc is treated as an independent orbit with the state vector approximately at the mid-arc; the state vector consists of a position vector and a velocity vector (i.e., six parameters per sub-arc) which are weighted according to their reliability. In addition to the state vector, the trajectory of each sub-arc depends on the potential coefficients. However, for sufficiently short sub-arcs, it is possible to enforce an a-priori (small) set of potential coefficients and greatly reduce the resulting errors in the satellite trajectory by slight adjustments of the state vector parameters.

The method used by AFGL leads to the formation of observation equations for satellite altimetry in which the parameter set is composed of the potential coefficients and of the state vector parameters. Owing to their highly patterned characteristics, the state vector parameters are eliminated from the total system and, if needed, later solved for efficiently. The exploitation of such patterns forms the backbone of the short arc mode of satellite altimetry. The important point is that as far as each satellite sub-arc is concerned, the potential coefficients are considered as fixed quantities rather than adjustable parameters, which plays a role in the computer savings mentioned earlier.

Used together with the gravity anomaly model in a simultaneous adjustment, the short arc mode of satellite altimetry has resulted in significant contributions to the determination of the global geoid and the earth's gravity field. When the past observational precision is considered, two recently described mathematical models associated with the global solution appear to be sufficient. For example, the standard error of a typical mean gravity anomaly representing a $5^0 \times 5^0$ geographic block is of the order of a few milligals, whereas the error entailed by various approximations in the gravity anomaly model amounts to just a few tenths of a milligal. Similarly, the observational precision of satellite altimetry achieved with GEOS-3 satellite has been typically in the vicinity of one meter; by comparison, the error attributable to the short arc property of holding the potential coefficients fixed within each sub-arc and adjusting its state vector parameters has amounted to a few decimeters. As another example, one may consider the error due to using the customary -- and efficient -- formula giving the geoid undulation in terms of the potential coefficients; this error hardly reaches more than a few centimeters or perhaps a few decimeters (but certainly less than 0.5m), depending on the geographic location.

However, the demands upon the mathematical models and the corresponding computer software necessarily increase as significant improvements in data gathering techniques take place. For example, the advances of SEASAT or a similar observational system over GEOS-3 capability to measure the sea surface topography amount to an order of magnitude in instrumental accuracy, perhaps half an order of magnitude in orbital accuracy(recently

represented by 10m to 20m in position), and an order of magnitude or more in terms of coverage density. When matched by development of more accurate algorithms as well as algorithms providing for an efficient adjustment of localized data, these advances will result in corresponding improvements in the knowledge of the gravity field and its fundamental surface, the geoid.

Following are the areas of research undertaken in this study whose intent is to make a contribution toward improving the determination of the earth's gravity field.

In a satellite altimetry adjustment, the radial distance r from the origin of the coordinate system (geocenter) to a subsatellite point on the "initial geoid" as defined by the initial potential coefficients is needed in conjunction with each altimetry measurement (such measurements may number in tens of thousands). An efficient way to compute this distance is to add the geoid undulation N as given by an approximate but simple formula in terms of the potential coefficients to the radial distance r' associated with the geocentric reference ellipsoid. However, the error in the approximation may reach a few decimeters, which under the new conditions becomes prohibitive. Another possibility is to compute r in an iterative process (it figures on both sides of the formula for r; on the right-hand side r appears in powers of a/r, fairly close to unity, where a is the earth's equatorial radius). To achieve the required accuracy, two iterations are needed even if the starting value is r', a very good approximation to r indeed. This approach would be, of course, costly in terms of computer run-time. In Chapter 2, an algorithm will be designed embodying the advantages of both methods while avoiding the disadvantages of either of them. It will serve to compute N -- and thus r -- to a centimeter accuracy without iterations.

As indicated previously, the short arc mode of satellite altimetry introduces an error of perhaps a few decimeters in final values of r, due to considering the potential coefficients constant within each sub-arc when forming the observation equations. Although the bulk of the total "raw" error (a few meters) in satellite positions is eventually accommodated by adjusting the state vector parameters, a further improvement in the results is necessary. However, since the errors (both the "raw" and the final errors) depend to a large extent on the length of an individual sub-arc, an improvement in accuracy may be achieved without modifying the short arc concept. An analysis concerned with the length of sub-arcs linked to the quality of the final results will be presented in Chapter 3.

Another improvement in accuracy will be achieved by refining the familiar gravity anomaly model. It has been pointed out that an error of a few tenths of a milligal may be introduced at certain geographic locations due to simplifications in this model. In the appendix and in Chapter 4, these simplifications will be traced down in order to eliminate the most important source of error. Although the errors entailed are small considering the precision of the data which have been used in actual computations, they should be minimized since they have the nature of systematic errors, at least regionally. Furthermore, they may become increasingly significant with higher data density and more accurate measurements.

The software development at AFGL has been progressing along two main avenues: 1) the determination of the earth's gravity field and the global geoid through the spherical harmonic potential coefficients, and 2) the determination of localized geoidal features achieved by various

methods. One of these methods used recently has been based on the notion of node point parameters. However, the situation which occurs most often in practice suits neither of the two strategies perfectly.

Although the globe as a whole may be covered reasonably well by different types of useful observations, there are invariably some areas with much denser coverage than others. For instance, the mean gravity anomalies of varying precision may be available for most of the earth's surface, but the satellite altimeter measurements may exist only in some areas such as the North Atlantic. Increasing the number of potential coefficients in the adjustment to resolve the required geoidal features in the North Atlantic would be wasteful since an exceedingly large number of coefficients would have to be added to determine the necessary detail. The core space of most computers would be inadequate for such a task. In addition, the problem would be ill-conditioned because of the global character of the potential coefficients (local data is insufficient for their determination). On the other hand, adopting the localized approach alone would preclude the utilization of the global type of data (here gravity anomalies or vast amounts of satellite altimeter data outside the area of interest); consequently, a global solution or global predictions of any kind would be impossible to obtain.

The objective of a more detailed, yet economical representation of the geoidal surface and certain features of the gravity field will be achieved by introducing point mass parameters into the model conceived in terms of potential coefficients. This will allow for a combination of both types of data (gravity anomalies, satellite altimetry) in both types

of observational distributions (global, local). A first, global solution will be carried out through an adjustment of a given set of spherical harmonic potential coefficients; a subsequent adjustment in terms of a suitable number of point mass parameters will serve to determine detailed geoidal features in the area of a high observational density. The final outcome of this method, described in Chapter 5, will be geoid undulations and/or gravity anomalies, as well as standard errors in these quantities, representing the local, along with the global, solution.

Most of the chapters in this report as well as its appendix are presented essentially in a self-contained manner. An advantage of this format is that they can be read independently without undue damage to a good understanding of the material. A typical example in this respect is Chapter 2 which has been also presented separately as [Blaha, 1978].

2. AN ACCURATE NON-ITERATIVE ALGORITHM FOR COMPUTING THE LENGTH OF A POSITION VECTOR TO A SUBSATELLITE POINT

2.1 Derivation

The radial distance (r) from the origin of the coordinate system (geocenter) to the geoid as defined by the spherical harmonic potential coefficients can typically be obtained in an iterative progress. This process was indicated e.g. in (2.2) of [Blaha, 1977] and is represented by (2.1) below which can be obtained in a straightforward manner from the formula for W (potential in the actual field), upon writing $W = W_0$ (potential on the geoid) and upon multiplying the whole equation by r/W_0. In practical applications related to satellite altimetry, the locations on the geoid for which r is sought correspond to subsatellite points. Their geocentric latitude and longitude are assumed known, derived from satellite altimetry.

Even if some crude initial value for r is adopted, e.g. the mean earth radius or an equatorial radius, a sub-millimeter accuracy is generally achieved on the third iteration. If this initial value is taken as the radial distance (r') to the reference ellipsoid such as the mean earth ellipsoid, the process is significantly accelerated; a sub-meter (in fact, a sub-half meter) accuracy is achieved on the first iteration and a sub-millimeter accuracy, on the second iteration. Although a sub-millimeter accuracy is clearly not needed, two iterations will be necessary in most cases since the first iteration results would be unsatisfactory even when using this

best available value r', associated with the rotating equipotential ellipsoidal model. This statement may be corroborated by pointing out that SEASAT-A altimeter capability is accredited with a 0.1 m standard error and that this (internal) accuracy should not be degraded by completely avoidable computational errors. Keeping in mind, then, the need for computational accuracy coupled with efficiency, our goal is to design an algorithm which would yield results characterized by a sub-decimeter or, perhaps, a sub-centimeter accuracy already from the first iteration.

The most important potential coefficient appearing in the formula for r is C_{20}, its value being approximately -0.00108. The next largest coefficients (i.e., some second, third, and fourth degree coefficients) are smaller than C_{20} by three or more orders of magnitude; their individual contributions to the value of r are smaller by at least two orders of magnitude when compared to the contribution of C_{20}. A similar statement holds true also with regard to the error in r as computed on the first iteration, caused by the approximation in the initial value. To eliminate the bulk of this error, the task at hand is to derive a correction term based only on the term containing C_{20} and on the rotation term (the contributions from these two terms have a comparable order of magnitude). Accomplishing this goal will mean that if, for instance, the first (uncorrected) iteration yields r to better than one meter, the error in the corrected result will be reduced to perhaps a sub-centimeter level with almost no additional computing effort.

The earlier mentioned formula for r reads as follows:

$$r = r_o \left[1 + \sum_{n=2}^{\infty} (a/r)^n \sum_{m=0}^{n} (C_{nm} \cos m\lambda + S_{nm} \sin m\lambda) P_{nm} (\sin\bar{\phi})\right]$$

$$+ \tfrac{1}{2} \omega^2 r_o r^3 \cos^2\bar{\phi}/(kM), \qquad (2.1)$$

where

$$r_o = kM/W_o \qquad (2.2)$$

and where W_o is the potential of the geoid; kM is the product of the gravitational constant and the earth's mass; a is the earth's equatorial radius; ω is the angular velocity of the earth's rotation; C_{nm} and S_{nm}, also referred to as C's and S's, are the conventional spherical harmonic coefficients of the earth's potential; $P_{nm}(\sin\bar{\phi})$ are the conventional Legendre functions in the argument $\sin\bar{\phi}$; and $\bar{\phi},\lambda$ are the geocentric latitude and longitude, respectively, of the (geoidal) point whose distance from the geocenter is being sought. In practical computations, the series expansion (2.1) is customarily truncated at some suitable n = N.

In considering the level ellipsoid that shares with the actual earth the values of ω and kM, from (2.1) we may deduce (the two centers of mass and axes of rotation coincide):

$$r' = r_o^* \left[1 + C_{20}^*(a/r')^2 P_2(\sin\bar{\phi}) + C_{40}^*(a/r')^4 P_4(\sin\bar{\phi}) + C_{60}^*(a/r')^6 P_6(\sin\bar{\phi})\right]$$

$$+ \tfrac{1}{2} \omega^2 r_o^* r'^3 \cos^2\bar{\phi}/(kM). \qquad (2.3)$$

The coefficients C_{20}^*, etc., are now related to the flattening of the ellipsoid, those beyond C_{60}^* being completely negligible; in the ellipsoidal field

we have $C^*_{20} = -J_2$, $C^*_{40} = -J_4$, etc., where the "J" coefficients appeared e.g. on page 73 of [Heiskanen and Moritz, 1967], from which the above formula can be readily verified (add the rotation term to obtain the normal potential U, specialize to U_o for the level ellipsoid, and multiply both sides by r'/U_o, where our r' was written as r in the above reference). Further,

$$r^*_o = kM/U_o, \tag{2.4}$$

where U_o is the potential of the ellipsoid. With minor changes in notations, (2.3) corresponds to the last equation on page 243 of [Blaha, 1977].

Let us now consider (2.1) where r on the right-hand side is replaced by r'. This gives rise to a correction term (c) which, in accordance with our earlier discussion, is the differential of the terms containing C_{20} and ω. With this provision, (2.1) can be written as

$$r = r_o(1+p) + \tfrac{1}{2}\omega^2 r_o r'^3 \cos^2\bar{\phi}/(kM) + c, \tag{2.5}$$

where

$$p = \sum_{n=2}^{\infty} (a/r')^n \sum_{m=0}^{n} (C_{nm}\cos m\lambda + S_{nm}\sin m\lambda)\, P_{nm}(\sin\bar{\phi}), \tag{2.6}$$

$$c = (r_o/r')\left[-2C_{20}(a/r')^2 P_2(\sin\bar{\phi}) + (3/2)\omega^2 r'^3 \cos^2\bar{\phi}/(kM)\right](r-r'). \tag{2.6'}$$

To compute c, we substitute $\frac{1}{2}(3\sin^2\bar{\phi} - 1)$ for $P_2(\sin\bar{\phi})$ and approximate r' by mean earth's radius ($\bar{R}$),

$$\bar{R} = 6371 \text{ km};$$

further, the Geodetic Reference System 1967 (GRS 1967) suggests the use of the following rounded values:

$$a = 6378.2 \text{ km},$$

$$r_0 = 6363.7 \text{ km},$$

$$\omega = 0.72921 \times 10^{-4} \text{ rad/sec},$$

$$kM = 3.9860 \times 10^{14} \text{ m}^3/\text{sec}^2.$$

This yields

$$c = \tilde{c}\,(r - r'), \tag{2.7}$$

where

$$\tilde{c} = c_1 \sin^2\bar{\phi} + c_2, \tag{2.8a}$$

$$c_1 = -3.0033\, C_{20} - 0.005169, \tag{2.8b}$$

$$c_2 = 1.0011\, C_{20} + 0.005169. \tag{2.8c}$$

If C_{20} is replaced by C_{20}^* in the GRS 1967, i.e., by -0.0010827, we have

$$c_1 = -0.001917, \tag{2.8b'}$$

$$c_2 = 0.004085. \tag{2.8c'}$$

Upon introducing the notation

$$\Delta r_0 = r_0 - r_0^*,$$

(2.5) is rewritten as

$$r = r_0^*\,(1 + p) + \tfrac{1}{2}\omega^2 r_0^* r'^3 \cos^2\bar{\phi}/(kM)$$

$$+ \Delta r_0\,(1 + p) + \tfrac{1}{2}\omega^2\,\Delta r_0 r'^3 \cos^2\bar{\phi}/(kM) + c. \tag{2.9}$$

The combination of (2.9) and (2.3) yields

$$r - r' = \tilde{N} + \Delta r_o(1 + p + t \cos^2\overline{\phi}) + c, \tag{2.10}$$

where

$$\tilde{N} = r_o \left[p - C^*_{20}(a/r')^2 P_2(\sin\overline{\phi}) - \ldots\right], \tag{2.11a}$$

$$t = \tfrac{1}{2}\omega^2 r'^3/(kM),$$

and where c was given in (2.7), (2.8). When dealing with t, we can again use the same numerical values that led to (2.7), giving

$$t = 0.001725. \tag{2.11b}$$

The value of r-r' needed for the evaluation of c from (2.7) can be safely replaced by $\tilde{N} + \Delta r_o$, as can be gathered from (2.10). Since Δr_o is usually zero (or a very small quantity), the error in c due to this approximation can reach at most 2 mm; this would occur when c itself reaches its maximum value of 0.49 m. From (2.10), we thus have

$$r - r' = \tilde{N}(1 + \tilde{c}) + \Delta r_o\ (1 + p + t \cos^2\overline{\phi} + \tilde{c}). \tag{2.12}$$

The value of geoid undulation (N) can be taken as

$$N = r - r', \tag{2.13}$$

whose maximum error was shown during numerical evaluation to be less than 0.7 mm (see [Blaha, 1977], page 60).

The formula (2.12) together with its variations will now be discussed. First of all, it should be noticed that r' figuring on the left-hand side as well as on the right-hand side in the expression for p-- and thus also for $\tilde{N}$-- does not have to be computed by (2.3) which has been

merely used in the theoretical derivation. Rather, a standard closed-form expression can be used with advantage for practical computations; it appeared e.g. in (4.23) of [Blaha, 1977] and it reads

$$r' = a/[1 + e^2 \sin^2\bar{\phi}/(1 - e^2)]^{\frac{1}{2}}. \tag{2.14}$$

Using the series expansion truncated at degree $n = N$, from (2.11a) together with (2.6) we obtain

$$\tilde{N} = r_o^* \sum_{n=2}^{N} (a/r')^n \sum_{m=0}^{n} (\Delta C_{nm} \cos m\lambda + \Delta S_{nm} \sin m\lambda)\, P_{nm}(\sin\bar{\phi}), \tag{2.15}$$

where

$$\begin{aligned} \Delta C_{20} &= C_{20} - C_{20}^*, \\ \Delta C_{40} &= C_{40} - C_{40}^*, \\ \Delta C_{60} &= C_{60} - C_{60}^*, \end{aligned} \tag{2.16}$$

the other ΔC's and ΔS's being essentially the C's and S's themselves. Even if $\Delta r_o = 0$, the above formula for $\tilde{N}$ will not yield, in general, the correct geoid undulation N. In particular, the difference between the two quantities amounts to $\tilde{c}\tilde{N}$ which could be as high as 0.49 m. Thus, in some applications the non-iterative result $\tilde{N}$ will not be satisfactory, but its corrected version will. We mention that in the spherical approximation, such that $r_o^* = a = r' = \bar{R}$, the above formula for $\tilde{N}$ reduces to, e.g., (27) of [Rapp, 1974] or (3.29) of [Needham, 1970], acceptable nevertheless for a variety of tasks (for instance, in determining the covariance function).

If $C_{20} = C_{20}^*$, the value of ΔC_{20} in (2.16) is zero; further, from

(2.8), (2.8') we have

$$\tilde{c} = -0.001917 \sin^2\bar{\phi} + 0.004085.$$

It then follows from (2.12) and (2.13), together with (2.11b), that

$$N = \tilde{N}(1 + 0.004085 - 0.001917 \sin^2\bar{\phi}) + \Delta r_o(1 + p + 0.005810 - 0.003642 \sin^2\bar{\phi}). \tag{2.17}$$

Most often in practice $r_o = r_o^*$ and thus $\Delta r_o = 0$, in which case (2.17) becomes

$$N = \tilde{N}(1.004085 - 0.001917 \sin^2\bar{\phi}). \tag{2.18}$$

The correction to $\tilde{N}$ has now a very simple form, namely

$$\tilde{c}\tilde{N} = (0.004085 - 0.001917 \sin^2\bar{\phi})\tilde{N}. \tag{2.19}$$

If $\tilde{N}$ were left uncorrected, the error in the geoid undulation for a given $\tilde{N}$ would be maximum on the equator, where it would reach $0.004085\,\tilde{N}$. When considering the largest magnitude of geoid undulations such as 120 m, this maximum error could amount to 0.490 m.

In conclusion, a non-iterative formula for N has been presented as (2.12). Adopting the GRS 1967 values for all the parameters except r_o, this formula reduces to (2.17). When also r_o is assumed to be given in this system, the final formula is (2.18). The non-iterative solution for r is simply

$$r = r' + N,$$

where r' has been given in (2.14). Computer simulations with a geopotential model truncated at degree and order (20,20) have indicated that in this way, an agreement with the three-iteration results obtained from (2.1) can be expected to within 0 - 2 mm in most cases (exceptionally, to within 4 mm). This amounts to important computer savings, considering that in real data reductions of satellite altimetry, the radial distance r is to be computed at thousands of locations.

2.2 Discussion

In order to explain the background of the proposed formulas from a slightly different angle, we adopt the standard simplification

$$U_o = W_o. \tag{2.20}$$

This means that the value of the normal potential on the level ellipsoid is equal to the value of the actual potential on the geoid, as used e.g. in [Heiskanen and Moritz, 1967], pages 83, 84. We have already defined the pertinent level ellipsoid as sharing ω and kM with the actual earth. If the ellipsoid shares with the actual earth also a fourth independent parameter such as the C^*_{20} coefficient, it is then the mean earth ellipsoid mentioned at the beginning. On the global scale, the value of r' computed for this ellipsoid via equation (2.14) represents the best available value to be used in the formula for r, i.e., on the right-hand side of (2.1).

Let us now rewrite (2.1) as follows:

$$r = F(r), \tag{2.21}$$

which indicates that r appears also on the right-hand side of the equation. Using the standard approach in which the advantage is taken of the rotating equipotential ellipsoidal model through the adoption of the best available value r' as described above, two choices can be made: 1) one can compute the radial distance to a subsatellite point from the first iteration and accept the result, in general good to better than half meter, namely

$$r_1 = F(r = r'); \tag{2.22}$$

or 2) one needs a more accurate result and proceeds to a second iteration (the improvement will be more pronounced than needed in practice),

$$r_2 = F(r = r_1). \tag{2.23}$$

By comparison, the main idea behind our method is to obtain r in (2.21) by a Taylor expansion, neglecting the second- and higher order terms:

$$r = F(r = r') + (dF/dr)_{r=r'} \, (r - r'). \tag{2.24}$$

First, the term $F(r = r')$ is identified. Since (2.20) implies that

$$r_o^* = r_o \tag{2.25a}$$

or, equivalently,

$$\Delta r_o = 0, \tag{2.25b}$$

from (2.1), (2.3), and (2.15) one can readily deduce

$$F(r = r') = r' + \tilde{N}. \tag{2.26}$$

This formula is equivalent to (2.22). The effort needed for computing $\tilde{N}$ is about the same (if not lighter) as that deployed in (2.22) in computing r_1. Comparing (2.24) and (2.26) with (2.12), where the second term is now zero due to (2.24b), one concludes that $(dF/dr)_{r=r'}$ must be our $\tilde{c}$, the factor $(r - r')$ in (2.24) being replaced by $\tilde{N}$ as in (2.12). In the following paragraph, we shall verify this explicitly since the "$\tilde{c}$-term" has been the backbone of the whole demonstration.

The replacement of r on the right-hand side of (2.1) by r' and the differentiation of the resulting expression with respect to r' yields

$$(dF/dr)_{r=r'} = (r_o/r')\Big[-\sum_{n=2}^{\infty} n(a/r')^n \sum_{m=0}^{n} (C_{nm}\cos m\lambda + S_{nm}\sin m\lambda)\, P_{nm}(\sin\bar{\phi}) + (3/2)\,\omega^2\, r'^3 \cos^2\bar{\phi}/(kM)\Big]. \qquad (2.27)$$

Numerically, the contribution of the term containing C_{20} in the above series is much greater than that of all the other terms with $n,m \neq 2,0$. Having a practical purpose in mind, we conclude that these terms can be safely neglected; the thus introduced error, the error caused by replacing $(r - r')$ by $\tilde{N}$ in the original expression $(dF/dr)_{r=r'}\,(r - r')$, and the error of neglecting higher order terms in (2.24) all amount to the total error of a few millimeters in N or r (this total error has been mentioned to exceptionally reach 4 mm). But the neglection of those terms in (2.27) implies that

$$(dF/dr)_{r=r'} \simeq (r_o/r')\Big[-2\,C_{20}\,(a/r')^2\, P_2(\sin\bar{\phi}) + (3/2)\,\omega^2\, r'^3 \cos^2\bar{\phi}/(kM)\Big] = \tilde{c}, \qquad (2.28)$$

upon considering (2.6') together with (2.7).

Collecting the results, we have

$$r = r' + \tilde{N} + \tilde{c}\tilde{N}, \qquad (2.29)$$

where r' is given in (2.14), $\tilde{N}$ in (2.15) with (2.25a) taken into account, and $\tilde{c}$ in (2.28), developed further in equations (2.8). We have seen that (2.29) is nothing else but the Taylor expansion (2.24). All the errors combined amount merely to a few millimeters. This of course follows from using the best available r' in the computations. With a different ellipsoid, the geoid undulation (N) as well as its approximation $\tilde{N}$ could be much greater than 120 m; thus the value $\tilde{c}\tilde{N}$ could be much larger than 0.49 m and the error caused by using it in lieu of $\tilde{c}(r - r')$ could be much larger than 2 mm reported

prior to (2.12). On ther other hand, if, with such an ellipsoid, $cN >> 0.49m$ as just stated, the one-iteration procedure leading to r_1 or $\tilde{N}$ could fail to result even in a sub-meter computational accuracy. In fact, it is precisely under such circumstances that the derived correction would be the most valuable; it would upgrade the poor accuracy of one-iteration results to, e.g., a centimeter accuracy.

3. FEASIBILITY OF THE SHORT ARC ADJUSTMENT MODEL IN SEASAT-A ALTIMETRY REDUCTIONS

3.1 Introduction

Traditionally, the satellite orbit determination and the ensuing geoid, ground stations, etc., determinations through various measurement techniques have been carried out using the mathematical model termed here the "long arc mode". Throughout the discussions, the data provided by the satellite will consist of altimeter data. For the present purpose, a long arc may be considered to span at least ¼ revolution. In theory, the long arc mode can approximate the "real world" of an orbiting satellite to any desired accuracy. One set of parameters in this mode are the dynamic parameters defining the forces acting upon the satellite. They include the earth's mass, the spherical harmonic potential coefficients C_{nm} and S_{nm} (they are denoted for the sake of brevity as C's and S's), the drag, the solar radiation pressure and other parameters. Since the dynamic parameters affect the satellite state continuously, so do the errors in these parameters, in the sense that their effect is cumulative. The longer the arc, the higher the number of dynamic parameters that have to be included in the adjustment in order to yield results good to a predetermined accuracy. It then follows that many hundreds of these parameters may be present in a long arc adjustment, entailing considerable computer core and run-time requirements, yet the errors beyond a certain arc length may be prohibitive.

There is no doubt that the long arc mode is essential for the establishment of satellite ephemeris and for other global tasks and applications. However, in certain applications great savings and superior results may be achieved by first using satellite ephemeris obtained through global tracking and then abandoning the long arc concept altogether in favor of what will be called the "short arc mode". In this mode, a long arc is divided into a number of short arcs of less than ¼ revolution and each such arc is treated as an independent orbit with the epoch at mid-arc. In a subsequent adjustment, the state vector of each such arc is subject to a-priori weighting consistent with the quality of the original reference orbit. This leads to a simultaneous recovery of all the state vector parameters and of the terrestrial parameters (in some applications it may be geoidal parameters such as C's and S's or certain node point values, etc., and in others it may be ground station coordinates, etc.).

The important advantage of the short arc mode in the geoid determination from satellite altimetry is that the reduction algorithm designed by Brown [1973] makes the computer run-time merely a linear function of the number of arcs. There is no limitation on the number of arcs which may be processed; the core space associated with the orbital part of the adjustment needs to provide only for one arc since it is re-used by every consecutive arc. This adjustment algorithm is described in detail, for example, in [Blaha, 1975], Chapter 2. An important characteristic of the short arc mode is that the C's and S's appearing in the orbit integration act merely as constants. On the other hand, the complete or partial geoid

may be represented by an adjustable set of C's and S's, by adjustable node point parameters, point masses, density layers, or combinations thereof. Although the described adjustment algorithm leads to important savings in the adjustment process, the first step toward its acceptability has to contain a demonstration that the desired accuracy of the final results could indeed be satisfied.

The accuracy analysis that first demonstrated the usefulness of the short arc mode appeared in [Brown, 1967]. That effort was intimately linked to the GEOS-3 satellite and its altimeter precision characterized by one-meter rms; the commensurate precision of 1 meter in the geoid determination was all that could be expected even under favorable conditions (sufficient data, no systematic errors, etc.). The use of the short arc mode in the case of GEOS-3 was justified by showing that orbital errors resulting from the enforcement of a reasonably accurate set of potential coefficients truncated at a fairly low degree and order can be accommodated by slight adjustments of the six state vector parameters. The residual modeling errors were shown to be inconsequential when compared to the overall precision of the altimeter, provided the arcs were sufficiently short.

A similar but more detailed analysis will be carried out in the following sections. It will be quite general as far as the description, mathematical formulation and some of the outcomes are concerned. However, the emphasis will be gradually shifted toward the GEOS-3 system and the SEASAT-A system which we shall then be able to investigate simultaneously. For this purpose, two kinds of ephemeris will be considered whose precision indicators (sigmas) are distinctly different. The coarse ephemeris,

supplied by NSWC after a preliminary long arc adjustment, are considered as having 10-20 m sigmas in position components and 0.1 m/sec - 0.2 m/sec sigmas in velocity components. They are roughly equivalent to the broadcast ephemeris associated with the Navy Navigational Satellite System. For this reason, they will be treated and identified as "broadcast ephemeris". On the other hand, the "precise ephemeris" will be characterized by approximately ten-fold improvements in precision over the coarse or "broadcast" ephemeris; this is also the relationship that exists between the precise and the broadcast ephemeris in the above mentioned NNSS.

The precision of the altimeter on board GEOS-3 has been described by a 1 m sigma, while the altimeter on board SEASAT-A is credited with a 0.1 m sigma. In the forthcoming discussion, 1 m sigma in altimetry will always be associated with the broadcast ephemeris and 0.1 m sigma in altimetry will always be associated with the precise ephemeris. The generated satellite orbit, however, will not always correspond to GEOS-3 or SEASAT-A.

In fact, the preliminary computer simulations will be carried out with a satellite whose approximately circular orbit is almost twice as high as that of GEOS-3 or SEASAT-A; its inclination is 60^0, its altitude about 1500 km and its period about 116 minutes, corresponding to 12.3 revolutions per day. A satellite with these specifications has been employed in the computer simulations described in Section 12.0 of [Brown, 1973]. The advantage of this "preliminary" satellite is that when it is used, the errors inherent in the short arc mode have smaller magnitude and more regular behavior than those encountered when dealing with the two lower orbiting

satellites. Thus, indications as to the feasibility of the geoid determination good to 0.1 m may be obtained at an early stage, and any further attempts may be discarded if negative results are produced under these favorable circumstances. Moreover, valuable knowledge, relating the short arc modeling errors to the state vector parameters, may be more easily gathered and discerned with a higher orbiting satellite. Finally, should such a satellite be launched at a later date, some results of the pertinent analysis could then prove useful in practice. From the academic point of view, an understanding of short arc adjustment results and characteristics is facilitated if more than one class of satellites are investigated.

At a later stage of computer simulations, the actual GEOS-3 and SEASAT-A satellites will be considered. Both have approximately circular orbits. The mean altitude of GEOS-3 is 843 km and its inclination is 115^0 (a symmetric case of 65^0 in inclination can equivalently be used in the analysis); its period is about 102 minutes, resulting in 14.1 revolutions per day. The altitude of SEASAT-A is given between 794 km and 808 km, its inclination is stipulated as 108^0 (or 72^0 in the symmetric case) and its period as 100.75 minutes, corresponding to 14.3 revolutions per day. The characteristics of the two satellites are so similar that for the purpose of this analysis, only one type of orbit, closely related to SEASAT-A, will be generated.

It should be emphasized that there exist other methods concerned with the geoid determination which make use of altimeter data from short satellite passes. Perhaps the most widely used technique in this respect

is the crossover adjustment as described in [Marsh et al., 1978]. Such an adjustment is used in practice in limited areas containing ascending and descending passes which generate a number of crossover points. The (interpolated) sea surface height differences at these points are minimized in a least squares process and the satellite passes are corrected for a bias and a tilt.

By comparison, the short arc mode leads to more general corrections than those for a bias and a tilt associated with the radial direction. In particular, the curvature of a satellite pass can also be changed, due to the possibility of adjusting the six state vector parameters. In addition to these parameters, the behavior of each arc depends on the gravity field, which fact is reflected in the orbital integration yielding satellite positions. The short arc mode is characterized by a simultaneous adjustment of all the altimeter observations, six state vector parameters per arc weighted according to their reliability, and the geoid parameters (usually completely free to adjust).

After this brief comparison, the reasons leading to the choice of the short arc mode over the crossover technique as the topic of the present analysis are listed below.

1) In the short arc mode, all altimeter observations on a given arc participate directly in the adjustment of that arc (their participation is not limited to a few measurements associated with the crossover points; besides their low number, such points may have the disadvantage of an unbalanced distribution along the arc).

2) In the short arc mode, the satellite passes are treated in line with the dynamic theory of motion allowing, among other things, for the curvature changes (an adjustment providing only for the bias and tilt corrections associated with the radial direction would result in a less rigorous treatment of satellite orbits; furthermore,since no orbital errors would be accommodated by curvature changes, unnecessary rejections of some passes could occur during the editing process).

3) In this mode, the reliability of the a-priori information is taken into account through appropriate weighting of the state vector parameters (the corrections to the orbital path do not depend exclusively on the ground misclosures).

Although the crossover technique certainly has practical merit due to its relative simplicity and efficiency, especially when compared to the long arc approach, it is believed that the short arc mode allows for a more rigorous treatment of the available information. Moreover, the efficiency of the adjustment algorithm concerned with the state vector parameters makes this mode economically no less advantageous than the crossover technique. By virtue of these considerations, the short arc mode is singled out as the fundamental method to be pursued in this study.

3.2 Mathematical Outline of the Short Arc Orbital Investigation

In order to see clearly the contribution of various error components in the short arc mode, it is instructive to describe the model at the level of observation equations. In accordance with Section 2.4 of [Blaha, 1977], reappearing on pp. 37-43 of [Blaha, 1977'], we express the model equation as

$$H = R - r + d, \tag{3.1}$$

where H represents the altimetry, R is the distance from the geocenter to the satellite at the time of the altimeter observation, r is the distance from the geocenter to a sub-satellite point on the sea surface, which has been assumed to coincide with the oceanic geoid, and d is a correction, always smaller than 5 m for the satellite altitude under 1000 km, which does not depend on any parameters under consideration.

In the above reference, R was assumed to be a function of only the state vector (s.v.) parameters; its value was obtained from the inertial coordinates computed by the orbital integrator with C's and S's acting merely as constants. Although this relationship will be preserved for adjustment purposes, in the error analysis we shall consider R to depend also on C's and S's. In view of the immediate need, we thus adopt the symbolic functional relationship

$$R \equiv R(\text{s.v.}; C,S). \tag{3.2a}$$

In earlier investigations, the value of r was considered to be a function of only the geoidal parameters, represented here by the adjustable C_{nm} and S_{nm} coefficients. However, in Section 2.4 of the above reference, a simple correction in the pertinent partial derivatives was introduced that expressed the dependence of r -- weak though it may be -- also on the state vector parameters. This correction was exceedingly simple and it made the adjustment model more rigorous without any penalty in terms of the computer run-time or storage requirements. Accordingly, we write

$$r \equiv r(C,S;\ s.v.). \tag{3.2b}$$

The earth's mass as well as the potential of the geoid are considered constant.

The usual linearization of the model equation (3.1) leads to the observation equation

$$V = [\partial(R-r)/\partial(s.v.)]d(s.v.) - [\partial r/\partial(C,S)]d(C,S) + [\partial R/\partial(C,S)]d(C,S) + L, \tag{3.3}$$

where

$$L = R(s.v.^b;\ C^0,S^0) - r(s.v.^b;\ C^0,S^0) + d - H^b, \tag{3.4}$$

and where

V = residual,

H^b = altimeter observation,

$s.v.^b$ = state vector parameters considered as observations (they are always weighted),

$d(s.v.)$ = corrections to $s.v.^b$ values,

C^0,S^0 = initial (approximate) values of C_{nm} and S_{nm},

$d(C,S)$ = corrections to C^0,S^0.

Suppose that the observations H^b in the observation equation (3.3) are errorless, such that

$$H^b = R(s.v.^b;\ C,S) - r(s.v.^b;\ C,S) + d, \tag{3.5}$$

where symbolically,

$$C,S = C^0,S^0 + d(C,S).$$

Disregarding the second order effects, (3.3) thus becomes

$$0 = -[\partial r/\partial(C,S)]\ d(C,S) + [\partial R/\partial(C,S)]\ d(C,S) + L \tag{3.6}$$

as can be immediately verified upon considering (3.4) and (3.5). The least squares adjustment would necessarily produce the same result as (3.6), where

$$V = 0, \quad d(s.v.) = 0,$$

because this corresponds to the (absolute) minimum of the total weighted sum of squares.

Let us now consider the model in which C's and S's in the orbital integrator are constants. Since they are subject to no corrections, this amounts to leaving the term $[\partial R/\partial(C,S)]d(C,S)$ out of the observation equation, whereby a modeling error is introduced in the short arc adjustment. The observation equation (3.3) is transformed into

$$\tilde{V} = [\partial(R-r)/\partial(s.v.)]\ \tilde{d}(s.v.) - [\partial r/\partial(C,S)]\ \tilde{d}(C,S) + L \tag{3.7}$$

where "~" indicates that in the new model, the values in question will possibly differ (after the least squares adjustment) from their counterparts in (3.3). It should be noted that the existing short arc adjustment technology, described in detail in Chapter 2 of [Blaha, 1975] and in Chapters

2, 3 and 4 of [Blaha, 1977] (Chapter 2 reappeared as [Blaha, 1977']) is represented by the observation equation (3.7) and not by the more rigorous observation equation (3.3); for example, (3.7) appeared in a similar form on page 26 of [Blaha, 1977] or, equivalently, on top of page 43 of [Blaha, 1977']. The above theoretical shortcoming has been more than offset in the past by the efficiency of the short arc mode. However, in view of the future requirements, the modeling error and its consequences should be scrutinized.

This task is facilitated by adopting the same errorless observations H^b as in (3.5). After the adjustment, the values of $\tilde{V}$ and $\tilde{d}$(s.v.) will directly represent the distortions in the residuals and in the state vector parameters, respectively, since in the correct model these values were zero (see equation 3.6). Similarly, the distortions in C's and S's will be $\tilde{d}(C,S) - d(C,S)$. The quantity responsible for all the distortions is of course the term left out of (3.3), whose value is $R(s.v.^b; C,S) - R(s.v.^b; C^0,S^0)$ and which has to be absorbed into the residuals (directly), the state vector parameters (indirectly through the model) and the geoidal parameters (indirectly through the model). If the latter two groups happened not to be affected at all, the entire discrepancy between the current and the correct adjustment model would be absorbed by the residuals. Each new residual would then become equal to minus the value of the neglected term (rather than zero), namely

$$\tilde{V} = \tilde{L},$$

where $\tilde{L}$, called the misclosure, is expressed as

$$\tilde{L} = R(s.v.^{b};\ C^{0},S^{0}) - R(s.v.^{b};\ C,S). \qquad (3.8)$$

Considering the systematic trends in the misclosures within individual arcs, such a contamination of the residuals would make their subsequent use for a detailed geoid representation virtually worthless.

Fortunately, as it has been already observed by Brown [1967] and as it will be shown in more detail later, the bulk of the misclosures can be accommodated by small changes in the state vector parameters if they are given appropriate freedom to adjust. It is then possible that the distortions in the adjusted geoid and in the residuals could become tolerable, consistent with the altimeter precision. Thus both the global geoid and its subsequent local detail could conform to reasonable accuracy requirements. However, the analysis related to the quality of the residuals will be presented in the next two sections. The purpose of the discussion below is to point out that under favorable circumstances, the potential coefficients are distorted very little, due to the fact that the dependence of R on C's and S's is much weaker than that of r, and that R and r depend on these coefficients each in a different manner.

If the satellite altitude were about 1500 km, the misclosures in short arcs of about 7.5 minutes in duration would under certain realistic conditions resemble those below (the intervals on either of the two arcs are 100 seconds, the units are meters, the state vectors are given at mid-arcs):

-.93, -.57, -.29, -.11, -.01, -.01, -.10, -.29, -.55, -.90; (3.9a)

+.29, +.18, +.09, +.03, .00, .00, +.03, +.09, +.15, +.23. (3.9b)

Near the mid-arcs the misclosures would be practically zero and on the ends they would reach maximum values of about 0.4 m or 0.5 m rms. At least one-third of each 7.5 minute arc and almost every entire arc shorter than 3 minutes would produce virtually zero misclosures and the remaining misclosures would vary in both magnitude and sign, depending on the location and on the direction of the satellite pass. If the satellite altitude were about 800 km, similar results would be produced upon shortening of all the time intervals by approximately 25%. Assuming a good global data coverage, the misclosures are expected to have a fairly random behavior introducing little or no bias in the geoid determination, although within one arc they are systematic. This can be explained by the fact that large scale systematic deformations of the geoid along one arc would be resisted by the observations along other arcs in that area as well as over the rest of the globe; the adjusted observations are tied together by the common parameters (adjustable C's and S's), whose number should allow for significant redundancy. The higher the redundancy, the better the chance that the effect of the misclosures will be neutralized, especially with respect to the global geoid. Helpful in strengthening the solution would be some additional observations of a different kind, having a good distribution over the entire globe. By virtue of these considerations, the misclosures are expected to be absorbed by the state vector parameters, the residuals, or both.

After a global adjustment via C's and S's, a detailed local adjustment may be carried out in terms of "local" parameters. For example, in [Blaha, 1977], Chapter 9, an adjustment was described in terms of the point mass parameters; it was based on the residuals obtained in the first, global adjustment. It then follows that if the residuals are not distorted to any significant extent, no difficulty should be experienced in representing faithfully the local detail to within the required resolution. One of the main tasks of this analysis, then, is to show if and under what conditions the distortions in the residuals can be inconsequential.

As a passing thought, we mention that caution would have to be exercised if one were dealing from the beginning with the geoid surface over a small region, without any previous global adjustment. In such a situation, bias in the local geoid could occur. Since there exist only few short arcs compared to the global case and since they all occur under somewhat similar circumstances (nearly the same location), the random character of the misclosures could be partially impaired. Furthermore, certain systematic deformations of the geoid in one region would not be resisted by observations over the rest of the globe. As a consequence, it is important in such cases to use arcs sufficiently short so that the misclosures inherent in the short arc mode may be almost entirely absorbed by very small adjustments in the state vector parameters.

Based on the discussion that followed equations (3.9), we expect that under favorable conditions (dense and reasonably regular global coverage, sufficiently short arcs, etc.), the geoidal parameters will not be

contaminated by the modeling error to any significant degree so that it can be assumed that

$$\tilde{d}(C,S) = d(C,S).$$

In the present context of errorless observations, this means that the only distortions to be considered are $\tilde{V}$ and $\tilde{d}(s.v.)$. Accordingly, (3.7) becomes

$$\tilde{V} = [\partial(R-r)/\partial(s.v.)]\tilde{d}(s.v.) - [\partial r/\partial(C,S)]d(C,S) + L. \qquad (3.10)$$

Upon replacing L by its value from (3.4) with the aid of (3.5), the last two terms are seen to equal $\tilde{L}$ and the observation equation may be written as

$$\tilde{V} = [\partial(R-r)/\partial(s.v.)]\tilde{d}(s.v.) + \tilde{L}. \qquad (3.11)$$

The analysis of $\tilde{V}$ and $\tilde{d}(s.v.)$ could be based on (3.11) where $\tilde{L}$ would be computed according to (3.8) and the partial derivatives $\partial(R-r)/\partial(s.v.)$ by the algorithm given in equation (2.22) or (4.27) of [Blaha, 1977]; the least squares adjustment would yield $\tilde{d}(s.v.)$ and hence $\tilde{V}$.

However, an equivalent adjustment can be carried out using the existing software based on the observation equation (3.7). In particular, an observation equation may be formed as follows:

$$\tilde{V} = [\partial(R-r)/\partial(s.v.)]\tilde{d}(s.v.) - [\partial r/\partial(C,S)]\Delta(C,S) + \{R[s.v.^b; (C,S) - d(C,S)] - r(s.v.^b; C,S) + d - H^b\}, \qquad (3.12)$$

where the terms within the braces equal $\tilde{L}$ given in (3.8). To see this, we only have to realize that the first term is $R(s.v.^b; C^0,S^0)$ and the second, third and fourth terms yield $R(s.v.^b; C,S)$ upon replacing H^b by its value

from (3.5). The arrangement of the terms within the braces indicates how $\tilde{L}$ is obtained without any modification of the existing algorithm which proceeds according to the formula (3.4). When such software is used for the present special task of adjusting (3.12), the initial values C^0,S^0 are replaced in (3.4) by the final values C,S. At this point, the feature needed for obtaining $\tilde{L}$ while "blindly" computing L is the shift of C,S by -d(C,S) to be applied only before the computation of R, i.e., before the use of the orbital integrator. Since this should be done for each satellite event, the procedure is facilitated by creating and storing two sets of C's and S's, one final and one shifted. The final set is heavily weighted, which results in nearly zero corrections $\Delta(C,S)$. But in this way, the second term on the right-hand side of (3.12) is zero and the adjustment yields the same $\tilde{V}$ and $\tilde{d}$(s.v.) as if (3.11) were utilized instead.

Another simple feature that can be used, this time in computing the residuals, is the following. In the above adjustment, $\tilde{d}$(s.v.) is computed for each arc. In a subsequent new adjustment, the state vector parameters are updated (as are C's and S's whose values, however, change insignificantly in this case) so that in (3.12), the quantity in braces is increased by $[\partial(R-r)/\partial(s.v.)]\tilde{d}(s.v.)$ as compared to its value in the previous adjustment. But this means that it has become $\tilde{V}$. The term "misclosures after the adjustment" is thus equivalent to the term "distortions in the residuals" (i.e., $\tilde{V}$), and they will be used interchangeably. The new adjustment has now fulfilled its role and may be stopped since the new corrections would be zero (good convergence properties are assumed). The two features just described are used in the analysis performed via computer

simulations which is the subject of the next two sections. However, apart from its practical usefulness, the observation equation (3.12) is of no further interest.

We shall next make two assumptions which will be verified later, and examine what they entail for the adjusted radial distances ($\tilde{R}$) from the geocenter to the satellite. The observations are again treated as errorless and the only distortions considered are those caused by the short arc modeling error. The first assumption states that the distortions in the residuals are very small, in the sense that the exhibit random behavior and their magnitude is in general smaller than the observational noise. The second assumption states that the positional part of $\tilde{d}$(s.v.) is nearly zero and that the bulk of the modeling error is accommodated through the velocity corrections alone. This assumption indicates that the absense of modeling error in R near the epoch is properly reflected in the corrections $\tilde{d}$(s.v.). In fact, a less restrictive assumption about these corrections would be sufficient. Inspired by the second assumption, we assert that in spite of the inherent modeling error in the short arc mode, the adjusted radial distances contain no significant distortions, no matter to which event on the arc they refer.

The mathematical proof of this assertion is based on the observation equation (3.10) and on the first assumption above. Combined with (3.4) in which H^b is given by (3.5), equation (3.10) becomes

$$\tilde{V} = [\partial R/\partial(s.v.)]\tilde{d}(s.v.) - [\partial r/\partial(s.v.)]\tilde{d}(s.v.) + R(s.v.^b; C^0,S^0) - R(s.v.^b; C,S). \qquad (3.13)$$

The short arc mode formula giving $\tilde{R}$ is

$$\tilde{R} = R[s.v.^{b} + \tilde{d}(s.v.); C^{0},S^{0}];$$

as we have seen, C^{0},S^{0} are not updated or otherwise modified, acting merely as constants in the orbital integrator. Equivalently, we write

$$\tilde{R} = R(s.v.^{b}; C^{0},S^{0}) + [\partial R/\partial(s.v.)]\tilde{d}(s.v.).$$

On the other hand, the corresponding errorless value is

$$R = R(s.v.^{b}; C,S).$$

Accordingly, the first and the third term on the right-hand side of (3.13) yield $\tilde{R}$, while the fourth term is R. Hence,

$$\tilde{R} = R + \tilde{V} + [\partial r/\partial(s.v.)]\tilde{d}(s.v.). \qquad (3.14)$$

From the analysis which will be described later, it follows that the distortions accommodated by the state vector parameters may reach 1-2 m in the case of the SEASAT-A system and 4-8 m in the case of somewhat longer arcs in the GEOS-3 system. In other words, the magnitude of $[\partial(R-r)/\partial(s.v.)]\tilde{d}(s.v.)$ may reach 1-2 m and 4-8 m, respectively. However, according to the statement on page 29 of [Blaha, 1977] or 44 of [Blaha, 1977'], the maximum contribution of the part $[\partial r/\partial(s.v.)]\tilde{d}(s.v.)$ is approximately 0.6% of the total contribution, here about 1 cm and less than 5 cm, respectively. The last term on the right-hand side of (3.14) is therefore completely negligible and we have

$$\tilde{R} = R + \tilde{V}. \qquad (3.15)$$

But this means that if the first assumption is justified, the radial distances to the satellite as obtained from the short arc adjustment can be expected to be undistorted by the modeling error.

3.3 Preliminary Analysis

The most important outcome of the introductory section has been the assertion that small corrections of the state vector parameters could possibly accommodate the discrepancies arising from the fixed potential coefficients in the short arc orbital integration. The obvious questions arise: 1) Will this possibility be translated into actual accommodations in altimetry reductions? 2) What are the conditions for this to occur? 3) What will be the effect of the leftover discrepancies on the residuals?

The investigation aimed at answering these questions began as a follow-up effort to that described in [Blaha, 1977]. Its purpose was to verify once again the already established methodology used in conjunction with the GEOS-3 satellite and to determine if the short arc technology could be made compatible with the new accuracy requirements for the SEASAT-A altimetry reductions. It was considered that the previous requirements could well be satisfied by the original short arc mode, but modifications and improvements might be needed in view of the tenfold more stringent requirements associated with the SEASAT-A system.

The most obvious modification that comes to mind is to provide for the partial derivatives of R with respect to all adjustable C's and S's and to form the observation equations as indicated in (3.3). However, in this way the efficiency of the short arc mode would be drastically impaired. Alternately, one could consider the formation of these derivatives only for a few most important coefficients. This procedure has not proved to be desirable for the following reasons. From an analysis of misclosures in the

observation equations (3.11), it appears that if the extended partial derivatives were to include a coefficient set as large as (4,4), the modeling error, although significantly reduced, would still be incompatible with even the one-meter precision ascribed to the GEOS-3 altimeter. It then follows that unless the arc is exceedingly short (e.g., 6 minutes in duration for 1 m accuracy and 2 minutes for 10 cm accuracy) this feature alone, without the accommodation by the state vector parameters, cannot correct the inherent defect of the presently used short arc technology. By comparison, computer simulations have indicated that slight changes in the state vector parameters could possibly accommodate the errors in R caused by the truncation beyond the degree and order (2,2) almost as well as it could accommodate the errors caused by the (4,4) or higher degree and order truncations, depending of course on the arc's length.

Another indication that the formation of the partial derivatives with respect to a small set of C's and S's would be of little help is the fact that often it is precisely such a small set of coefficients that is kept unchanged in the adjustment. For instance, the values of the set (2,2) would usually be completely fixed, while the remaining values in the set (6,6) could be weighted according to their reliability, etc. We mention that if the geoid were not described through the adjustable C's and S's, the present considerations of the partial derivatives would be altogether meaningless. The most promising approach in either case, then, appears to be the use of the existing short arc technology whenever it is justified. This amounts to reverting to the three questions asked earlier and attempting to provide answers with the help of computer simulations.

In order to study the misclosures $\tilde{L}$ shown in (3.11), a "new" or errorless set of reasonable C_{nm} and S_{nm} coefficients is chosen, including the degree and order (8,8). Errorless state vector parameters are generated in the long arc mode upon utilizing these coefficients in the orbital integrator. The satellite orbits simulated in this section are characterized by 1500 km altitude according to the description given earlier. For this purpose, the orbital integrator is furnished the six initial state vector components (in inertial coordinates) as follows:

$x = 7{,}894{,}450\,m, \qquad \dot{x} = 0,$

$y = 0, \qquad \dot{y} = 3{,}552.25\,m/sec,$

$z = 0, \qquad \dot{z} = 6{,}152.93\,m/sec.$

A number of "old" sets of C's and S's have been created from the new set by truncations at a chosen degree and order and/or by shifting certain coefficients by 25% or 50% of their values toward zero. The "old" or shifted coefficients are used exclusively in the short arc computations and not in the geoid representation (the geoid is described by the "new" C's and S's) so that the constant terms in the observation equations consist of $\tilde{L}$ without contamination from any other error source. This corresponds to the numerical procedure described following (3.12); the errorless state vector parameters are $s.v.^{b}$, the "new" coefficients are (C,S), and the "old" coefficients are $(C,S) - d(C,S)$, where $d(C,S)$ are the above mentioned shifts. As an illustration, misclosures (in meters) from two 15 minute arcs are listed below. Each arc contains 10 events in 100 second intervals and its epoch is exactly at mid-arc (halfway between the fifth and sixth events);

the "old" set of C's and S's is represented by the coefficients truncated beyond the degree and order (2,2).

(3.16a)

+8.40, +5.16, +2.64, +.96, +.12, +.08, +.84, +2.20, +4.00, +6.08;

(3.16b)

-2.84, -1.84, -1.04, -.40, -.04, -.04, -.48, -1.32, -2.60, -4.24.

If the suppressed coefficients were instead shifted toward zero by only 25% of their values, the above misclosures would be 25% of those shown. It has been observed that such a linear relationship exists for various "old" sets not only before an adjustment, but also with respect to $\tilde{V}$, i.e., the misclosures after the adjustment, as well as to the corrections to the state vector parameters (and even for the extremely small corrections to the "new" C's and S's weighted heavily at their errorless values). Another linear property of the short arcs comes to light when shifting, for example, all the 3rd degree coefficients to zero and, in a different simulation, all the 4th degree coefficients to zero, etc.; when these shifts are applied simultaneously to produce another "old" set, the misclosures are the algebraic sums of the appropriate misclosures in the previous simulations, and the same holds true also after the adjustment.

A more instructive example than the one above follows. The arcs involved are half as long as those given previously (i.e., 7.5 minutes in duration as opposed to 15 minutes), containing again 10 events each with the epoch at mid-arc, but the interval is now 50 seconds. The C's and S's within (3,0) to (4,0) are shifted by 25% of their values toward zero, those within (5,0) to (6,6) are shifted by 50% and the remaining coefficients are shifted to zero (i.e., truncated). The misclosures for two out of many investigated

arcs are (they already appeared in 3.9a, 3.9b):

$$-.93, -.57, -.29, -.11, -.01, -.01, -.10, -.29, -.55, -.90; \qquad (3.17a)$$

$$+.29, +.18, +.09, +.03, .00, .00, +.03, +.09, +.15, +.23. \qquad (3.17b)$$

Although their behavior patterns are quite representative, the magnitude of the misclosures in (3.17a) is much higher than the average. One notices that the above misclosures can be well fitted by a segment of a parabola. Unlike in the previous example, the misclosures can now be almost perfectly compensated for by very small changes in the state vector. In particular, the three position components of d(s.v.) given by the adjustment are practically zero and, considering the average absolute values, the radial velocity component is about 1/100 or 1/1000 of the velocity sigmas (the first fraction applies in conjunction with the precise ephemeris and the second, with the broadcast ephemeris), while the horizontal velocity components are about 1/4 or 1/40 of the velocity sigmas. Such changes reduce the misclosures in (3.17a), (3.17b) to a 0-1 cm level. In the same series of experiments, the largest $\tilde{V}$ amounts to 4 cm. Even without adjusting the state vector parameters in these experiments, arcs of up to 2 or 3 minutes in duration would be compatible with the 0.1 m altimeter precision, and arcs of up to 5 or 6 minutes would be compatible with the 1 m precision.

These estimates are believed to be on the conservative side insofar as the "preliminary" satellite is concerned, since in the real data adjustment, the coefficients would be unlikely to change by the large amounts considered. In fact, the coefficients up to (4,4) would probably be fixed or weighted so that they might not change at all and further coefficients up to (6,6) could also be weighted. It is also believed that the overall

picture would change little if the analysis of the modeling error were to include higher degree and order coefficients than (8,8); the computer simulations have indicated that the contribution to the misclosures decreases quite rapidly with the increasing degree, especially for shorter arcs (under 10 minutes).

It has been noticed that if the arcs such as those represented by (3.17a), (3.17b) are extended to 15 or more minutes in duration, the misclosures lose much of the symmetry and they cannot be accommodated nearly as well as in shorter arcs, even if all the state vector parameters should change the best possible way. Some of the results obtained with 15 minute arcs will be reflected in the closing paragraphs of this section. In general, that part of the modeling error which cannot be absorbed by the state vector parameters is expected to propagate directly into the residuals. In the present context, it gives rise to the values $\tilde{V}$. The distortions in the residuals that can still be tolerated depend on the accuracy requirements; in principle, they should behave fairly randomly and their magnitude should be smaller than the observational noise. This is of course the important first assumption of the previous section, which led to the property of little or no distortion in radial distances to the satellite.

During the adjustments of simulated errorless observations in the short arc mode, C's and S's describing the geoid have been heavily weighted at their errorless values, the a-priori sigmas being 0.0001 of their absolute values except for C_{20} for which the factor 0.00001 has been utilized instead. The state vector parameters have been weighted according to two

fundamentally different schemes distinguishing between the broadcast and the precise ephemeris. The broadcast ephemeris may be conservatively described by the following sigmas:

20m, 20m, 20m, 0.2m/sec, 0.2m/sec, 0.2m/sec; (3.18a)

these values are associated with the in-track, cross-track, and radial directions in position (first three numbers) and in velocity (last three numbers). The positional sigmas used in [Blaha, 1977], page 32, are 24m, 17m, and 8m; however, the replacement of the above 20m by these values entails virtually no changes in the adjustment of the state vector parameters so that for the present purpose, such a distinction would be merely academic. On the other hand, the velocity sigmas used in this reference were approximately tenfold smaller than those in (3.18a).

Realistic estimates of the velocity sigmas for the broadcast ephemeris have been gathered from [Arur, 1977]. The counterparts of (3.18a) as recommended in this reference are

19m, 15m, 9m, 0.2m/sec, 0.1m/sec, 0.2m/sec. (3.18b)

No differences in the simulated adjustments have been observed (other than in a-posteriori sigmas of the state vector parameters) upon interchanging some or all of the corresponding values between (3.18a), (3.18b). With the altimeter sigma of 1m and the state vector sigmas taken from (3.18a) or (3.18b), the adjustments in the state vector parameters lead to almost perfect compensation of the misclosures as reported following equations (3.17).

An interesting observation can be made at this stage. If the freedom to adjust the velocity components of the state vector were restricted by lowering their sigmas to 0.02m/sec, the full accommodation of the misclosures by the state vector parameters in an otherwise unchanged problem would be inhibited. This points to the possibility of improving the final results by slightly relaxing the velocity sigmas (perhaps twofold) if they happened to be constrained too tightly. As an example, the velocity state vector components and their sigmas could be computed by a curve-fitting technique, based on the coordinates of many points along a satellite pass. Although the latter (positional) information would be supplied by the broadcast ephemeris, the sigmas of the computed velocity components could be significantly lower than the 0.2m/sec associated with the "raw" velocity information; for example, 0.05m/sec sigmas have been used in some practical applications. This number is believed to be quite suitable in general, in that it provides the freedom needed for the best possible accommodation of the modeling error, but in some cases it might prove beneficial to relax it to perhaps 0.1m/sec.

The values adopted during the second part of the computer simulations with the "preliminary" satellite are 0.1m for the altimeter sigma and

2m, 2m, 2m, 0.02m/sec, 0.02m/sec, 0.02m/sec (3.19)

for the state vector sigmas, everything else remaining unchanged. The remark made in the last paragraph about an occasional slight relaxation of the velocity components applies here as well, although all the pertinent values are decreased tenfold. The most surprising experience learned during these experiments is that whether using the broadcast ephemeris sigmas

such as given in (3.18a) or (3.18b) together with 1m altimeter sigma (hence called 1m observational system), or using the precise ephemeris sigmas given in (3.19) together with 0.1m altimeter sigma (hence called 0.1m observational system), the misclosures after an adjustment as well as the corrections to the state vector parameters are almost identical. The a-posteriori sigmas of the state vector parameters are of course different, while the misclosures before an adjustment are identical by definition. This finding enables us to analyze the needs and to answer some of the questions pertaining to either observational system with the help of the same computer simulations. From the computational point of view, then, the distinction between the two observational systems resides merely in the magnitude of misclosures, before and after an orbital adjustment, which can still be tolerated. As expected, the crucial limitation in either system, although to a different degree, is the arc length. The observational density on a given arc does not induce any improvement in the misclosures before or after an adjustment. In fact, when altimeter data rate was doubled, no significant changes could be detected in the pertinent results other than improved a-posteriori sigmas.

We may conclude this section by presenting an aggregated outcome of several experiments. If the anticipated variations in C's and S's were as large as those associated with equations (3.17), the expected accommodation of the misclosures by the state vector parameters would be excellent even in view of the 0.1m observational system, provided the orbital arcs do not exceed about 8 minutes in duration. A longer satellite pass could be divided into two or more independent arcs of acceptable length. In the

case of 7.5 minute arcs, the largest misclosure after an adjustment has been noted as 4cm. On the other hand, if these arcs were extended to 15 or 16 minutes, the residuals would be distorted by errors of up to 0.3m that could not be absorbed by the state vector parameters. However, considering the magnitude and the fairly random behavior of these distortions, such arcs would still be acceptable for the 1m observational system. The expected relatively small distortions $\tilde{V}$ in one or the other system confirm the validity of the "first assumption" encountered in the previous section. This outcome is not modified in the following section, other than by the necessity to shorten the pertinent arcs by small amounts.

Guided by these experiments, we can reach a plausible answer to the three questions asked at the beginning of this section. If the arc length is kept below approximately 8 minutes (1/14 revolution) in the case of the 0.1m observational system and below 16 minutes (1/7 revolution) in the case of the 1m observational system, the state vector parameters will tend to compensate for the inherent short arc modeling error in the sense that the left-over errors, absorbed mainly by the residuals, will likely be contained within the observational noise. Moreover, the satellite positions may improve, especially in the radial direction, via the revised estimates of the state vector parameters which include the altimetry information. In these considerations, the favorable circumstances such as good data distribution, absence of systematic errors, etc., are assumed throughout.

3.4 Realistic GEOS-3 and SEASAT-A Analysis

Following the suggestion from the introductory section, a satellite orbit, closely related to SEASAT-A, will be generated. It will serve for the analysis of the short arc modeling error in either the SEASAT-A system or the GEOS-3 system. The former is identified by the 0.1m observational system in conjunction with the present, "realistic" satellite, and the latter by the 1m observational system in conjunction with the "realistic" satellite. The orbit will be generated as follows. First, a nominal Keplerian orbit will be computed which approximates the desired orbit. Subsequently, the six initial conditions of the Keplerian orbit will be used in the gravity field represented by the set of C's and S's previously called "new" or errorless. Except for the different numerical values of the six initial conditions, the analysis will be completely equivalent to that of the last section.

The nominal Keplerian orbit is stipulated to be circular ($e = 0$), centered at the origin, the radius (r) being

$$r = a = 7{,}172\text{km},$$

where "a" is the semi-major axis of the Keplerian "ellipse". This number represents the earth's mean radius ($\bar{R} = 6{,}371$km) augmented by the mean satellite altitude ($\bar{h} = 801$km). The earth is considered a point mass located at the coordinate origin; this together with the previous stipulation is consistent with Kepler's first law. The orbit's inclination is specified to be

$$i = 72^0.$$

The product of the gravitational constant (k) and the earth's mass (M) is adopted as

$$kM = 3.986030 \times 10^{14}\ m^3/sec^2.$$

Applying the condition $e = 0$ to Kepler's second law, we obtain for the satellite's velocity in an inertial system:

$$ds/dt = (kM/a)^{\frac{1}{2}} = 7{,}455.04 m/sec.$$

From Kepler's third law, the mean motion in our case is computed as

$$n = (1/a)(kM/a)^{\frac{1}{2}} = 0.0010395 \text{ radians per second;}$$

accordingly, the period is

$$p = 2\pi/n = 100.74 \text{ minutes,}$$

which corresponds to 14.3 revolutions per day.

The satellite's line of nodes is chosen to coincide with the x-axis of the inertial coordinate system and at the time $t = 0$, the satellite is chosen to be located at the ascending node. The velocity component $\dot{x}$ at this location is zero, while the other two components are computed as

$$\dot{y} = (ds/dt) \cos i,$$
$$\dot{z} = (ds/dt) \sin i.$$

Thus at $t = 0$, the values of the six state vector components in inertial coordinates are

$$x = 7{,}172{,}000m, \qquad \dot{x} = 0,$$
$$y = 0, \qquad \dot{y} = 2{,}303.73m/sec,$$
$$z = 0, \qquad \dot{z} = 7{,}090.17m/sec.$$

When these values are supplied to the orbital integrator together with the (8,8) set of "new" C's and S's, the SEASAT-A specifications are recovered very satisfactorily. For example, the period is found to be 100.65 minutes, the highest geocentric latitude is listed as 71.98^0, and the highest and lowest altitudes above the generated geoid are listed as 807km and 786km, respectively. The subsequent equator crossings occur at intervals of about 25.4^0 in longitude.

The general findings of the previous section are confirmed also with the new satellite. Thus, essentially the same results (except for a-posteriori sigmas) are obtained whether the SEASAT-A system or the GEOS-3 system is involved in adjusting an otherwise identical problem, although the former represents approximately a tenfold improvement over the latter in both the altimeter and ephemeris precision. This property reduces the total number of computer simulations in the analysis of the short arc modeling error. It is again noticed that the observational density on a given arc does not lead to any significant changes in the misclosures before or after adjustment (a-posteriori sigmas of the state vectors are, of course, improved by higher observational density, especially in the radial direction). Consequently, there is no need to analyze SEASAT-A and GEOS-3 adjustment properties by different means, although the altimeter data rate attributed to SEASAT-A is much higher than that of its predecessors. Finally, the previous finding pertaining to an occasional relaxation of the stage vector velocity components applies also in this section. In particular, if for any reason the input velocity sigmas were much lower than the usual values associated with the 1m or 0.1m observational systems, slight relaxations

leading to perhaps 0.1m/sec or 0.01m/sec, respectively, could again prove beneficial for the accommodation of the modeling error by the state vector parameters.

In order to demonstrate the increased importance of the modeling error in the case of a lower orbiting satellite, we list below the misclosures for two 15 minute arcs roughly corresponding to (3.16a), (3.16b). The time intervals are exactly the same as those used previously and the "old" set of C's and S's is again the "errorless" set truncated beyond the degree and order (2,2).

(3.20a)

+13.27, +7.77, +3.74, +1.24, +.12, +.11, +.82, +1.88, +2.95, +3.83;

(3.20b)

-7.23, -4.43, -2.31, -.87, -.10, -.11, -1.03, -3.01, -6.09, -10.20.

These misclosures have a much larger magnitude than their counterparts in equations (3.16) and they are less symmetric about the epoch. The resulting values $\tilde{V}$ can be represented by a sinusoidal curve along the arc as follows:

+1.18, -.25, -.98, -1.00, -.46, +.32, +.97, +1.11, +.42, -1.31;

+.32, -.05, -.27, -.30, -.16, +.10, +.31, +.33, +.08, -.36.

Under these circumstances, the usefulness of 15 minute arcs would be marginal even for the GEOS-3 system.

We shall henceforth deal with a more realistic situation, where the C's and S's within (3,0) to (4,4) are shifted by 25% of their values toward zero, those within (5,0) to (6,6) are shifted by 50% and the remaining coefficients are truncated, just as described prior to equations

(3.17). First, arcs of almost 16 minutes in duration (950 seconds) are briefly considered in view of the GEOS-3 system. Each arc contains 20 events in 50 second intervals and its epoch is again exactly at mid-arc (halfway between the tenth and eleventh events). Although its misclosures are higher than the average, a typical arc is depicted as

(3.21)

+4.89, +3.95, +3.09, +2.32, +1.65, +1.09, +.65, +.33, +.11, +.01,
+.01, +.10, +.27, +.51, +.81, +1.15, +1.52, +1.91, +2.31, +2.72.

The distortions in the residuals result again in a sinusoidal curve, although with a somewhat lower amplitude than seen previously. For this arc, the distortions $\tilde{V}$ are

(3.22)

+.27, +.15, +.04, -.07, -.15, -.20, -.21, -.20, -.16, -.09,
-.01, +.08, +.16, +.22, +.25, +.24, +.17, +.05, -.15, -.42.

In some cases, however, the values of $\tilde{V}$ can reach 0.8m which, compared with the previous example, is not an improvement important enough to make the usefulness of such arcs much better than marginal.

We continue with an intermediate example which differs from the one above only by the arc length; it is now 7.5 minutes (450 seconds), resulting in 10 events at 50 second intervals on each arc. The two arcs listed below are chosen such that the first has larger misclosures than the average (here both before and after the adjustment), while the second is quite typical for this series of experiments. The misclosures are

(3.23a)

-1.81, -1.06, -.52, -.18, -.02, -.02, -.14, -.36, -.63, -.94;

(3.23b)

-.88, -.57, -.31, -.12, -.01, -.01, -.13, -.36, -.72, -1.19.

The values of $\tilde{L}$ on the first arc show a lack of symmetry, which causes greater distortions in the residuals. The corresponding values of $\tilde{V}$ are computed as

(3.24a)

-.12, +.03, +.09, +.09, +.05, -.02, -.08, -.10, -.04, +.12;

(3.24b)

+.05, -.02, -.04, -.03, .00, +.02, +.04, -.03, .00, -.04.

Their periodicity is still apparent, but their magnitude is greatly reduced.

When comparing the results (3.24) with (3.22) including the statement that followed, we notice the rapidity with which the values of $\tilde{V}$ decrease with shorter arcs. As far as the GEOS-3 system is concerned, arcs of 14 minutes in duration would most likely be acceptable. On the other hand, the results in equations (3.24) point to the possibility that even a small reduction in duration (from 7.5 minutes to, e.g., 6 minutes) could produce results acceptable for the SEASAT-A system. This possibility will be explored next.

The duration of arcs in this last series of experiments is slightly below 6 minutes (i.e., 350 seconds). Each arc contains 8 events at 50 second intervals, the epoch being at mid-arc (halfway between the fourth and fifth events). Of the two arcs listed, the first has again larger misclosures than the average and the second is again as typical as possible. The misclosures are

$$-1.25, -.63, -.22, -.02, -.02, -.19, -.49, -.90; \qquad (3.25a)$$

$$+.57, +.30, +.11, +.01, +.01, +.11, +.28, +.51. \qquad (3.25b)$$

We observe that the misclosures are quite symmetric, especially in (3.25b); as expected, this is reflected by smaller distortions in the residuals. The values of $\tilde{V}$ for these two arcs are computed as

$$-.05, +.02, +.05, +.03, \ .00, -.04, -.04, +.03; \qquad (3.26a)$$

$$+.01, \ .00, -.01, -.01, -.01, +.01, +.01, \ .00. \qquad (3.26b)$$

These results illustrate that arcs of 6 minutes (or less) in duration are indeed acceptable for the satellite altimetry reductions in the SEASAT-A system.

3.5 Conclusion

The experiments described in the last section have identified the short arc mode as a feasible adjustment method for satellite altimetry data reductions in both the GEOS-3 system and the SEASAT-A system. The former is associated with a 1m altimeter sigma and the broadcast ephemeris characterized by 10-20m sigmas in position, while the latter is attributed a 0.1m altimeter sigma and is envisioned to be used in conjunction with the precise ephemeris represented by 1-2m sigmas in position. The acceptability of the short arc mode in either system is based on the notion that the distortions in the residuals, caused by the inherent modeling error in the adjustment algorithm, should be contained well within the observational noise of the system considered. The reason rendering the short arc mode potentially capable of producing results of the required accuracy is that although the algorithm introduces a modeling error, it can, under suitable conditions, compensate for this error almost entirely through small adjustments of the state vector components (in particular, the velocity components).

The suitable conditions just mentioned involve to a great extent what the length of a satellite path may be in order to still be considered a "short arc" in one or the other system. Based on the analysis via computer simulations, the following criteria are suggested with regard to the applicability of the short arc mode. If altimetry data reductions are performed in the GEOS-3 system, the length of the arcs should be kept at or below 14 minutes in duration (1/7 revolution). If the SEASAT-A system is envisioned instead, this length should be kept at or below 6 minutes (1/17 revolution).

Finally, the reasons making the short arc methodology promising for SEASAT-A altimetry data reductions are recapitulated.

1) The ultra-precise reference orbits are not required. In fact, state vector positions good to within 2 meters appear to be completely satisfactory. (It is believed that the long arc approach would require about 10 or 20-fold improved ephemeris.)

2) The computer core and run-time requirements are modest (compared to the long arc approach, enormous savings can be realized).

3) In addition to the geoid determination, the simultaneous adjustment of the potential coefficients and the state vector parameters yields improved satellite positions. The added strength is supplied through the altimeter observations and is thus most noticeable in the radial direction.

4. ACCURACY IMPROVEMENT IN ADJUSTING GRAVITY ANOMALIES

The conventional gravity anomaly model has been evaluated and analyzed in Chapter 5 of [Blaha,1977]. The basic formula in this report, abbreviated here as [B], appeared on page 77 as

$$\Delta\tilde{g} = -\partial T/\partial r - (2/r)T \ ;$$

the symbol "~" indicates the presence of certain approximations, r is the radial distance from the coordinate origin to the point under consideration and T is the disturbing potential at the same point. When applied to the geoid, this formula expresses the fundamental boundary condition as can be gathered from [Heiskanen and Moritz,1967], page 88.

The formula for $\Delta\tilde{g}$ was shown in [B] to contain two approximations resulting in two kinds of errors, Δ_1 and Δ_2, called respectively "direction error" and "spherical error". Both these errors are negligible in various practical applications, being usually much smaller than 0.2 mgal. However, it is desirable to reduce their effect in certain theoretical considerations, especially with regard to a possible future development. It was shown in [B] that the spherical error is usually much larger than the direction error, and that its elimination could be an easy matter. The remaining direction error would then have the form (see page 78 of [B])

$$\Delta_1 = (g - g_r) - (\gamma - \gamma_r) \equiv g - \gamma - (g_r - \gamma_r) \ , \qquad (4.1)$$

and the expression for gravity anomaly would read

$$\Delta g \equiv g - \gamma = g_r - \gamma_r + \Delta_1 \ ; \tag{4.2}$$

g_r would be expressed in terms of spherical harmonic potential coefficients and γ_r would be computed very efficiently from

$$\gamma_r = \gamma - f(a, J_2, \ldots \bar{\phi}) \ , \tag{4.3}$$

that is, as a function of ellipsoidal parameters and the geocentric latitude (see Appendix 2 of [B]).

Corresponding to (4.2), the observation equation containing Δg or g (referred to the geoid) as observations is developed in the appendix of the present report. With the knowledge that only the direction error is present, we omit writing the symbol Δ_1 explicitly. This observation equation, presented in (A.47) of the appendix, is written as

$$v = C \{-F(1/r_o)dr_o + \sum_{n=2}^{N} [(n+1) - F(r_o/r)] (a/r)^n \, dS(n)\} + g_r^0 - g_r \ , \tag{4.4}$$

where, from (A.51b),

$$g_r = \Delta g + \gamma - f(a, J_2, \ldots, \bar{\phi}) \ , \tag{4.5}$$

as can also be gathered from (4.2), (4.3); if the observations were represented by g, $\Delta g + \gamma$ in (4.5) would simply be replaced by g. The other quantities are described as follows:

$$\gamma = (a \, \gamma_e \cos^2\phi + b \, \gamma_p \sin^2\phi)/(a^2 \cos^2\phi + b^2 \sin^2\phi)^{\frac{1}{2}} \ ,$$

Somigliana's formula where a,b are the semi-major and semi-minor axes, respectively, of the reference ellipsoid, γ_e, γ_p are the values of normal gravity at the equator and the poles, respectively, and ϕ is the geodetic latitude;

$$F = (2 + P_4 + D)/[1 + (r_o/r)(P_2 - 3D/2)];$$

$$g_r^o = C(1 + P_3 - D) ;$$

$$C = kM/r^2 ,$$

$$D = (\omega^2 r/C) \cos^2\bar{\phi} ,$$

$$P_2 = \sum_{n=2}^{N} n(a/r)^n S(n) ,$$

$$P_3 = \sum_{n=2}^{N} (n+1)(a/r)^n S(n) ,$$

$$P_4 = \sum_{n=2}^{N} (n+1)(n+2)(a/r)^n S(n) ,$$

$$S(n) = \sum_{m=0}^{n} (C_{nm} \cos m\lambda + S_{nm} \sin m\lambda) P_{nm}(\sin\bar{\phi}) ,$$

$$dS(n) = \sum_{m=0}^{n} (dC_{nm} \cos m\lambda + dS_{nm} \sin m\lambda) P_{nm}(\sin\bar{\phi}) ;$$

the quantities r, r_o, W_o, kM, ω, C_{nm}, S_{nm}, $P_{nm}(\sin\bar{\phi})$, $\bar{\phi}$, λ have been presented in (2.1), (2.2) and in the text which followed. All the coefficients in the observation equation (4.4) are computed with the initial values of the parameters (these values are subsequently updated in a least squares adjustment).

The radial distance r from the geocenter to the point associated with a gravity-type input value can be computed by the efficient algorithm of Chapter 2, yielding a sub-centimeter accuracy without iterations. In

conjunction with each such point -- as well as with each subsatellite point in a satellite altimetry adjustment -- the values of sin $m\lambda$ and cos $m\lambda$ are needed for m varying between 0 and n. Since numerical evaluations of these functions must be performed a great many times (depending on the data set), the need for a maximum efficiency in the adjustment algorithm calls further for the adoption of the well-known recursive formulas:

$$\cos m\lambda = 2 \cos\lambda \cos(m-1)\lambda - \cos(m-2)\lambda \ ,$$
$$\sin m\lambda = 2 \cos\lambda \sin(m-1)\lambda - \sin(m-2)\lambda \ .$$

Accordingly, the standard trigonometric series expressions are used only for $m = 1$; when $m = 0$ the problem is trivial, while for $m = 2$ the above recursive formulas already apply, reducing to

$$\cos 2\lambda = \cos^2\lambda - \sin^2\lambda \ ,$$
$$\sin 2\lambda = 2 \sin\lambda \cos\lambda \ .$$

We now present the observation equation (4.4) using the matrix notations. From (A.48), (A.49) developed in the appendix, we have

$$v = ax + \ell \ , \tag{4.6}$$

where

$$a = C \ \{-F(1/r_o); \ \ldots q(n) \ P_n(\sin\bar{\phi})\ldots, \ \ldots q(n) \cos m\lambda \ P_{nm}(\sin\bar{\phi})\ldots,$$
$$\ldots q(n) \sin m\lambda \ P_{nm}(\sin\bar{\phi})\ldots\} \ , \tag{4.7a}$$

$$q(n) \equiv \ [(n+1) - F(r_o/r)] \ (a/r)^n \ ; \tag{4.7a'}$$

$$x = [dr_o;\ \ldots dC_{no}\ldots,\ \ldots dC_{nm}\ldots,\ \ldots dS_{nm}\ldots]^T\ ; \qquad (4.7b)$$

$$\ell = g_r^0 - g_r \equiv C(1 + P_3 - D) - g_r\ , \qquad (4.7c)$$

with g_r given in (4.5) and described immediately below that equation.

5. GEOPOTENTIAL MODEL CONTAINING SPHERICAL HARMONICS AND POINT MASSES

5.1 Introduction

On a global scale, the geoid and some other quantities related to the gravity field (e.g. gravity anomalies) have been customarily represented in terms of the spherical harmonic expansion. The kinds of data often dealt with in physical geodesy are geoid undulations (the separation between the reference ellipsoid and the geoid) and gravity anomalies (the difference between the gravity reduced to the geoid and the normal gravity referring to the reference ellipsoid). The latter are usually given in the format of mean gravity anomalies representing certain blocks rather than individual points. If undulation and/or gravity anomaly data is supplied for the whole globe, the spherical harmonic potential coefficients can be solved for in a least squares adjustment. On the other hand, if a reliable set of these coefficients is adopted, geoid undulations, gravity anomalies (or, equivalently, gravity values referring to the geoid surface), and other quantities -- for example deflections of the vertical -- can be computed for the whole globe or for a specific region.

In order to represent the geoid to a desired resolution, the necessary number of potential coefficients may be exceedingly large. For example, the resolution corresponding roughly to 1^0 (about 111 km) would require a set of these coefficients complete through the degree and order (180,180), which means that almost 33,000 coefficients would be needed. To produce the resolution twice as fine (i.e., corresponding to about 0.5^0), four times as

many coefficients would be required, etc. If enormous quantities of data were available all over the globe so that such sets of coefficients would be theoretically solvable in a least squares adjustment, the matrix of normal equations in the first case above would have the dimensions of almost 33,000 x 33,000. The difficulties connected with the actual solution of this system would be formidable (there is no need for going into detail). To proceed in this way in order to describe the geoid over a relatively small region, such as the North Atlantic, would be extremely wasteful if at all possible.

In theory, an adjustment would not have to be performed if homogeneous data of uniform quality were available over the entire globe; in such a case, closed integral formulas could be employed. However, neither of these requirements are satisfied by the presently available data -- and will not be in the foreseeable future. It is therefore clear that the spherical harmonic approach offers only a partial answer when a detailed representation of the geoid over a limited area is contemplated.

One method which has been recently studied makes use of the local parameters called "point masses". Other technical terms describing this type of parameters have also been used in the geodetic literature, for example "buried masses" or "mass concentrations". The point masses could be used as quantities generating the disturbing potential from which all the geophysically meaningful quantities such as geoid undulations, gravity anomalies, etc., can be derived.

The disturbing potential mentioned above, added to the normal potential, results in the actual potential. A method where point masses give rise to the (total) disturbing potential has been presented in [Reilly et al.,

1978]. Another possibility is to separate the total disturbing potential (T) into two parts, $T_{S.H.}$ and $T_{P.M.}$; the first is represented by a spherical harmonic expansion and the second, by the point mass model.

The idea of supplementing the spherical harmonic expansion of the earth's potential by point masses is relatively recent, pioneered among others by Needham [1970]. The purpose of introducing the point mass parameters into an adjustment of gravity anomalies, satellite altimetry and perhaps other quantities as well is to add fine structure to a geopotential model based on the spherical harmonic coefficients. The point mass locations (including their depth, in theory stipulated with respect to the geoid or a reference ellipsoid) should be the object of a judicial choice, while their magnitudes are subjected to an adjustment in a generally over-determined system.

A study concerned with both the global and the local representation of the gravity field was reported in [Blaha, 1977], abbreviated here as [B]. In it, the global features of the gravity field were described via an adjustment of a set of spherical harmonic potential coefficients, and the detailed local features via an adjustment of point mass parameters. The first adjustment was conceived so as to accommodate the available data of geoid undulations (obtained from satellite altimetry) supplemented by mean gravity anomaly data. The second, localized adjustment was based on the residuals computed in a region of interest from the first adjustment; these residuals were to be accommodated in an additional data fit through the use of a suitable number of new parameters, the point masses. Thus the second adjustment was envisioned as superimposed locally on an adjustment of a set of potential coefficients.

This philosophy allows a relatively small total number of parameters to express such local detail of the gravity field which would otherwise (i.e., in the first type of adjustment performed alone) require perhaps a hundred-fold increase in the number of parameters. It has been suggested in the same reference that the results from the second adjustment, such as geoid undulations and/or gravity anomalies and their variances, could be added algebraically to their counterparts obtained from the first adjustment. The final result would then be contour maps of geoid undulations and/or gravity anomalies, constructed from the values computed in a grid which would extend over the region of interest. No point masses would be present beyond that region so that the geoid there would essentially coincide with the global geoid. Similar statements could be of course made with respect to the contour maps of standard deviations (sigmas) of the adjusted and/or predicted quantities.

The number of point masses in such an adjustment is a function of the required detail which, in turn, must be linked to the data distribution. It is believed that a reasonable strategy is to use one point mass per two data points in each direction, provided the data are fairly uniformly distributed in a certain area; this means that ideally there would be about four times as many data points as point masses. This strategy will be adhered to later on, in the analysis of computer simulations, although in practice its variations could be considered in cases where the data do not have the desirable distribution. Along with each observation entering the adjustment, a realistic sigma would have to be supplied.

The "required detail" is defined, in terms of the geoidal features, by the shortest half wavelength which should be taken into account by the

adjustment. In the angular units this half wavelength is symbolized by $\Delta\theta$ and in the linear units, by $s = R\Delta\theta$, where R is the earth's mean radius. The horizontal separation of point masses is denoted by s'. The point masses are assumed to be distributed, as well as possible, in an equilateral grid over the area of interest, the main practical reason being the need for a reasonably uniform representation of the geoid. This property would amount, in a certain sense, to a local equivalent of the spherical harmonic expansion resulting in the same resolution everywhere, regardless of the actual magnitude of the geoidal features or of the varying data density; in particular, if more data are available in one area, they are still smoothed (or filtered) to produce about the same detail as in the neighboring regions. From the theoretical point of view, the point masses should be located at approximately the same depth beneath the geoidal or ellipsoidal surface (in the spherical approximation, on a sphere) and a given depth-side ratio should be maintained. Both these properties were described in [Needham, 1970], abbreviated here as [N], and were recapitulated on pages 142, 143 and in Section 9.3 of [B]; the most desirable numerical value of the depth-side ratio was furthermore obtained under two stipulations discussed in these references, and was listed as 0.8. Clearly, an irregular distribution of the point masses would imply that at least one of the above two properties would have to be forsaken.

Two basic designs will be pursued in this study with regard to the density of regularly distributed point masses in an area of interest. The first, "original design", is characterized by

$$s' = s, \tag{5.1}$$

and the second, "dense design", is characterized by

$$s' = \tfrac{1}{2}s\,, \tag{5.2}$$

where s' is the grid size of the point masses. In accordance with a previous statement, the observations will be assumed to be given in an equilateral grid having $\frac{1}{2}s'$ on the side; this idealized case yields four times as many observations as parameters (regardless of the design considered -- original or dense). A useful variation of this arrangement is to use a geographic instead of an equilateral grid, resulting in about the same observational grid near the equator but not elsewhere, and to decrease the weights entering a least squares adjustment by the factor $\cos\phi$, where ϕ is the latitude of the observation point considered.

The study of point masses performed in [B] was concerned exclusively with the original design. Thus the ratio

$$d/s = 0.8\,, \tag{5.3}$$

where d represents the depth of the point masses, now has one of the following two meanings:

$$d/s' = 0.8 \ldots \text{original design,}$$

$$d/s' = 1.6 \ldots \text{dense design.}$$

The formula (5.3) is perhaps the most useful of the three because it gives the depth directly in terms of the required resolution. From the computer simulations analyzed in Section 5.4 it will become apparent that

$$d = 0.8s \tag{5.3'}$$

is indeed the most advantageous depth from the accuracy and other standpoints, and the various depths utilized will be indicated in terms of this "d".

The concept studied in both [N] and [B] is that of constraints associated with the point mass model. This concept can perhaps best be introduced by describing certain difficulties which may be encountered in practical computations. Since the gravity anomaly is obtained in part from the first derivative (in the radial direction) of the disturbing potential, it is a less smooth function than the geoid undulation which is essentially a linear function of the disturbing potential (in the spherical approximation it is exactly such a linear function). The geoid undulation data have been experienced to produce reasonably behaving gravity anomalies without any constraints imposed on the point mass adjustment while the converse may not be always true. In particular, a set of constraints may be needed which will guarantee that the large-scale geoidal features, as determined from the given potential coefficients, are undisturbed by the adjustment of gravity anomalies, and thus, that the detail described via point masses merely results in certain fluctuations from this "mean" surface. This could be accomplished with the aid of constraints specifying that the mean undulation over a given block is to be unaffected by the inclusion of point masses. The kinds and number of constraints, the size of the constraint blocks, etc., will be dealt with in the following sections.

A statement pertaining to the orientation and limitations of this study is in order. Since the most important data source considered in this study is the satellite altimeter, the main purpose of adjusting spherical

harmonic coefficients and point masses is to provide a fairly detailed description of the geoid, consistent with the data density and distribution. The mean gravity anomalies have been envisioned to serve only in the first adjustment performed in terms of spherical harmonics. Due to the data gaps over the continental regions, such an adjustment could not be performed with the use of satellite altimetry alone. In the second adjustment, however, the input provided by satellite altimetry would completely override any influence of the gravity anomaly input for two reasons; over the region of interest, such as the North Atlantic, the altimeter data density would be hundreds or thousands of times higher than the density of data coming from the other source and, furthermore, the weight associated with each altimeter observation would be much higher than the weight associated with any given gravity anomaly. Accordingly, the parameters in the second adjustment providing the local geoidal detail are the point mass magnitudes, and the "observations" in the area of concern are the minus residuals of geoid undulations as obtained in the first adjustment; the input weights are unchanged from those in the first adjustment.

Due to the local purpose and characteristics of the point mass parameters, they have not been envisioned to be used in computing satellite orbit perturbations. In fact, such a refinement of satellite orbits is not needed if the short arc adjustment model is utilized in the process of geoid determinations (or in other tasks such as Doppler positioning, etc.). If the orbital arcs in this model are chosen sufficiently short, no significant systematic errors are introduced in the geoid determination, in spite of the fact that only a relatively small set of spherical harmonic potential coefficients

is used in the orbital integrator and of the fact that this set is not subjected to any type of adjustment (these coefficients act merely as constants). The reason for this is that the modeling error due to the limited number of these coefficients and thus also to the absence of point masses or other equivalent parameters, etc., can be accommodated by slight adjustments of the state vector parameters. More detail on this subject is given in Chapter 3 of this report.

Initially, the point mass model was envisioned such that the spacing of point masses should correspond to the shortest half wavelength which it is desired to take into account, and such that the number of constraints should be approximately equal, in the global case, to the number of the spherical harmonic potential coefficients solved for in the first adjustment. Since one constraint may be considered as eliminating one parameter, this procedure (two adjustments) would involve about the same total number of parameters as in the case where only the first adjustment would be carried out with a larger set of potential coefficients representing the gravity field to the same half wavelength. In Section 5.3, the above number of constraints will be arrived at through more detailed considerations. Although these conclusions may seem obvious, they have to be approached with caution.

The main reason for caution is the realization that the point mass (P.M.) model does not possess a useful characteristic of the spherical harmonic (S.H.) model, in particular, that there are no inherent "orthogonality relations" in this model. By contrast, a S.H. adjustment of uniformly distributed undulation or gravity data of equal precision and sufficient density yields the same final values of the potential coefficients, geoidal heights,

etc ., regardless of whether all the coefficients are adjusted simultaneously or the coefficients through the degree and order (6,6), for example, are adjusted first, followed by the adjustment of the coefficients (7,0) through (8,8), etc., based on the previously computed residuals. All the coefficients thus obtained will be uncorrelated (due to the orthogonality relations). If "errorless" data is generated by a given set complete through the degree and order (12,12), for example, this set will be recovered exactly and the final residuals will be zero. However, if the data do not have the ideal distribution mentioned above, the adjusted potential coefficients will no longer be independent. Such a situation would lead to different sets of adjusted coefficients obtained via different routes and the same would hold for the adjusted observations. None of the adjusted coefficient sets would reproduce the original set and none of the residual sets would contain only zeros. The residuals would be expected to get larger with more intermediate stages in the total adjustment; by forcing separate adjustments, certain correlations would be suppressed, which would exercise an influence similar to an application of constraints (thus the rms residual would increase).

If the spherical harmonic coefficients were adjusted simultaneously with the point masses, the "errorless" data could not be reproduced even if the distribution were again ideal. The reason for this is the correlations which would now exist between the adjusted S.H. coefficients and the P.M. parameters, as well as among the latter parameters themselves. The constraints whose number would be equal to the number of coefficients present in the adjustment in accordance with an earlier statement would only make matters worse. If two adjustments were now performed instead of one in the

usual manner (the constraints would enter the second, point mass adjustment), the rms residual would further increase. Thus, an adjustment of the S.H. coefficients followed by an adjustment of the P.M. parameters with an appropriate number of constraints could not reproduce the "errorless" data exactly, although one, two, or more adjustments exclusively in terms of the S.H. coefficients could. The total number of parameters in both approaches would be about the same, but the correlations in the first approach would constitute an insurmountable obstacle to obtaining zero residuals. The first approach coincides in fact with the "original design" with constraints. This has lead to the conception of the "dense design" and to the reduction in the number of constraints -- if not their outright elimination in some cases -- as demonstrated in Section 5.4. In this way, the desired accuracy in the geoid determination could finally be approached.

5.2 Mathematical Background

In the spherical approximation, the boundary condition and Brun's formula, respectively, read

$$\Delta g = -(\partial T/\partial r)_{r=R} - (2/R)T \ , \tag{5.4}$$

$$N = T/G \ , \tag{5.5}$$

where

Δg = gravity anomaly,

N = geoid undulation,

T = disturbing potential,

R = earth's mean radius (6371 km),

G = average value of gravity (980 gals).

These two formulas can be found, for example, on pages 88 and 94 of [Heiskanen and Moritz, 1967]. They also appeared on page 184 of [B].

Although the above two formulas refer usually to the normal (or ellipsoidal) field, in [B] they referred to a higher order field, in particular, to that described by the S.H. model after the first adjustment. Thus, T in this reference depicted the disturbing potential associated with point masses and was intended to describe the field beyond the S.H. model (for example, the 14,14 model). Accordingly, the quantities Δg, N and T appearing in (5.4), (5,5) could be imagined with the subscript "tot" if they refer

to the normal field, and with the subscript "P.M." if they refer to a higher order field represented by the S.H. model. Since it holds by definition that

$$W = U + T_{tot} ,$$

where

W = actual potential,

U = normal potential,

T_{tot} = (total) disturbing potential,

and since we have chosen to separate the total disturbing potential into two parts,

$$T_{tot} = T_{S.H.} + T_{P.M.} , \tag{5.6}$$

it follows from (5.4) and (5.5) that

$$\Delta g_{tot} = \Delta g_{S.H.} + \Delta g_{P.M.} , \tag{5.7}$$

$$N_{tot} = N_{S.H.} + N_{P.M.} . \tag{5.8}$$

The parts denoted "S.H." can be treated either in the spherical approximation or in a more accurate form. An accurate formula for T appeared, for example, on page 80 of [B]. Sufficiently accurate expressions for N and Δg are presented in Chapter 2 and in the appendix of this report, respectively. On the other hand, the "P.M." part can always be treated in the spherical approximation; after the first, S.H. adjustment, the residuals to be accommodated by the subsequent P.M. adjustment are usually at least one or two orders of magnitude smaller than the original observations (e.g., geoid undulations) so that this kind of approximation could not

appreciably contaminate the resulting geoid representation, gravity anomalies, etc. By contrast, if the "total" quantities in (5.7), (5.8) were represented by a model containing the spherical approximation, the error thus introduced could amount to several decimeters in geoidal heights. It is thus clear that if the P.M. model containing the spherical approximation were used in an adjustment of the "total" quantities referring directly to the normal field, a decimeter accuracy in the geoid determination could not even be approached, due to this approximation alone.

It was stated earlier that even with an ideal data distribution and coverage, etc., a simultaneous adjustment of the S.H. and P.M. parameters would not yield the same results as the two separate adjustments considered in this study. A similar statement could be made with respect to the variances of the adjusted and/or predicted quantities. However, due to the obvious practical reasons we shall nevertheless add algebraically the corresponding quantities obtained in the two separate adjustments, in the manner of equations (5.7), (5.8). The same will hold true also for the corresponding variances.

Since we shall henceforth deal only with the "P.M." part of the quantities in (5.6) - (5.8), etc., the subscript "P.M." can be safely omitted. Based on the relation

$$T_i = \sum_j (1/\ell_{ij})(kM)_j \ , \qquad (5.9)$$

where

ℓ_{ij} = distance between the i-th (observation) point and the j-th point mass,

$(kM)_j$ = j-th scaled point mass,

from the formulas (5.4), (5.5) we can derive

$$\Delta g_i = (1/R) \sum_j (1/\ell_{ij})[(R^2 - F_{ij})/\ell_{ij}^2 - 2] (kM)_j , \tag{5.10}$$

$$N_i = (1/G) \sum_j (1/\ell_{ij})(kM)_j , \tag{5.11}$$

where

$$\ell_{ij} = (R^2 + R_1^2 - 2F_{ij})^{\frac{1}{2}} ,$$

$$R_1 = R - d ,$$

$$F_{ij} = 2RR_1 \cos \psi_{ij} ,$$

$$\cos \psi_{ij} = \sin\phi_i \sin\phi_j + \cos\phi_i \cos\phi_j \cos(\lambda_i - \lambda_j) .$$

The spherical approximation becomes further apparent from the fact that the point masses are considered to be located on a sphere of radius R_1 (at the depth d beneath the surface of the sphere approximating the earth).

The above equations represent an adapted version of the formulas which appeared in [N] and [B]. They also agree with the algorithm presented in [Reilly et al., 1978]. In this reference, however, the formulas similar to those above referred to the ellipsoid and not to a higher order surface; with the exception of the normal gravity, they were again based entirely on the spherical approximation.

The constraints encountered in [N] and [B] are the volume and the mass constraints; we have

$$\int_{\Delta A} N\,dA = 0 \quad \text{or} \quad \int_{\Delta A} T\,dA = 0 \ ,$$

which can be developed into

$$\int_{\Delta A} \sum_{j} (1/\ell_{ij})(kM)_j \, dA_i \ \ldots \text{ volume constraint} \ , \tag{5.12}$$

where

ℓ_{ij} = distance between the element dA_i and the point mass $(kM)_j$,

dA_i = surface element of ΔA,

ΔA = constraint block.

The computationally manageable form of (5.12) is described on page 202 of [B]. Further we have

$$\sum_{j} (kM)_j = 0 \ \ldots \text{ mass constraint} \ , \tag{5.13}$$

where j ranges over all the point masses within the block ΔA. The size of such blocks is one of the subjects of the next section.

5.3 Number of Parameters in an Idealized Point Mass Model

If the P.M. model had the outstanding property of the S.H. model as characterized by the orthogonality relations, an adjustment in terms of point masses would be expected to yield, under certain conditions, about the same results as a S.H. adjustment. This would also hold true for a combined adjustment of the S.H. and P.M. parameters, or for a P.M. adjustment based on a previous S.H. adjustment which is of most concern to us. The conditions just mentioned would imply:

a) a global type of adjustment compatible with the S.H. model (both the observations and the point masses would cover the entire globe),

b) an ideal distribution of observations (the observations would have sufficient density, uniform distribution and equal precision),

c) an ideal distribution of point masses (the point masses would be located in an equilateral grid having the length s on the side), and

d) the same number of parameters in both types of adjustment (if both types of parameters were equally "powerful", their numbers should be compatible so as to faithfully represent the same geoidal detail).

If an equivalence between the two types of adjustments could indeed be confirmed, the point masses could then be applied in a limited area (rather than globally), i.e., most of them could be eliminated, yet the local geoid representation would be preserved. We would thus be in the presence of a detailed local solution which would be as reliable as a corresponding

S.H. solution, except that the number of parameters to be solved for in a least squares adjustment would be drastically reduced.

If a required resolution, in terms of spherical harmonics, corresponds to the use of an (n,n) S.H. model, s is given approximately as $R\Delta\theta$, where (in degrees)

$$\Delta\theta^0 = 180^0/n \ . \qquad (5.14)$$

It turns out that the number of point masses distributed in a grid with s on the side is approximately equal to the number of parameters in the above S.H. model, i.e., approximately equal to $(n+1)^2$; a small number of S.H. coefficients held fixed at a given value ($C_{10} = C_{11} = S_{11} = 0$, etc.) and the equivalent number of relations which would be used in the P.M. model to define the coordinate system are inconsequential in this context. It thus follows that in an idealized P.M. model used alone, no constraints would be needed in order to produce about the same final resolution as that obtained through the corresponding S.H. model.

However, if the above idealized P.M. model were combined with the S.H. model developed through the degree and order $(\overline{n},\overline{n})$, the previous number of parameters would increase; the number of P.M. parameters would be unchanged but we would have additional $(\overline{n}+1)^2$ parameters present, whether the adjustment itself is split into the usual two adjustments or not. In this situation, the idealized P.M. model would therefore have to contain about $(\overline{n}+1)^2$ appropriate constraints, whose effect would be to reduce the total equivalent number of parameters to the previous level, in order to produce results similar to those obtained with the "purely" S.H. model (n,n) with no point masses.

In this manner, the equivalent number of parameters in the global case would be made essentially invariant with respect to the chosen degree and order of the spherical harmonic part of the complete model. It would then be also equal to the number of parameters in the "purely" S.H. model which would produce about the same final resolution. The kinds of constraints (volume constraints, mass constraints) and their mathematical formulation were described in the preceding section. In [N] and [B] their form was essentially the same as here, but the sizes of the blocks ΔA in which they were applied -- and thus their total number -- were quite different.

Considering the relation (5.14) and the text which followed, it becomes clear that if ΔA blocks had $R\Delta\bar{\theta}$ on the side, where

$$\Delta\bar{\theta}^0 = 180^0/\bar{n} , \qquad (5.15)$$

the number of these blocks would be approximately $(\bar{n}+1)^2$. This was also the number of constraints arrived at in [B] through heuristic reasoning. Thus, on the average, about one constraint in a $\Delta\bar{\theta} \times \Delta\bar{\theta}$ block (in angular units) would be required, where $\Delta\bar{\theta}$ corresponds to the half wavelength of the underlying spherical harmonic resolution. On the other hand, the number of constraints suggested in [N] was two per $\Delta\bar{\theta} \times \Delta\bar{\theta}$ block (one volume constraint and one mass constraint), i.e., the double of the number needed to achieve an equivalence in the number of parameters when comparing the different types of global adjustment.

The actual arrangement of constraints in [B] was the following. Instead of $\Delta\bar{\theta} \times \Delta\bar{\theta}$ blocks, it was suggested that the blocks should be defined as

$$\Delta A = 2\Delta\bar{\theta} \times 2\Delta\bar{\theta} \ , \tag{5.16}$$

and that two constraints (one volume constraint and one mass constraint) should be applied in each such block. In order to increase the number of these constraints two-fold so as to correspond, on the average, to one constraint per $\Delta\bar{\theta} \times \Delta\bar{\theta}$ block, the use of overlapping blocks of the same size was further suggested.

It was considered that if the validity of this arrangement could be confirmed, a very similar approach could be undertaken in practice, in cases where a high observational density exists only over a limited area A while the rest of the globe is covered by observations of lower density. One can imagine mean gravity anomalies as representing the lower density observations and satellite altimetry as giving rise to the high density observations extending, for example, over the North Atlantic region. Following the first (S.H.) adjustment, the point masses would be introduced only in the area A. The constraints whose form and mode of application have been described would also be used only in the area A.

Unfortunately, the actual P.M. model departs from the idealized model to such an extent that the accuracy of the results would be severely impaired if the guidelines conceived for the idealized case were followed in practical computations. The reason is, of course, the high degree of correlation among the P.M. parameters (if adjusted simultaneously, the

S.H. and P.M. parameters would also exhibit significant correlations), even in cases of an ideal global distribution of observations. The two immediate consequences of this consideration are:

a) the number of constraints suggested appears too high, not only in [N], but also in [B] (in some cases presented in the next section the best results will be seen as obtained with very few or no constraints), and

b) the number of P.M. parameters (distributed in a grid with s on the side) is too low to represent the geoid to a desired resolution, even in the absence of constraints.

In the accuracy analysis to be described in the next section, the constraints will be applied without modification, but cases with no constraints or with only one half of the constraints, i.e., with either mass constraints alone or volume constraints alone, will also be examined. This scrutiny of the effect of constraints will not change when additional point masses are utilized in the manner of the "dense design" mentioned earlier. The depth of the point masses will not be affected by this design, either. The analysis will be carried out exactly as in the case of the "original design", except that the point masses will be densified (two-fold) along each dimension of the grid.

5.4 Accuracy Analysis of the Point Mass Model

The key element in the present accuracy analysis is the presence of an "errorless" standard to which the results of individual adjustment solutions can be compared. This standard consists of geoid undulations or gravity anomalies generated through a S.H. model truncated at a certain degree and order. No random errors are attached to these values which represent the observations to be adjusted in two steps, using a S.H. model of a lower degree and order than the above, followed by an adjustment of a P.M. model. The latter model is intended to represent the features of the gravity field to the same resolution that characterized the generated data. Thus the comparison of the final, adjusted results with the generated values is meaningful, yielding the rms value of the fit which can serve in the making of a realistic assessment of the whole adjustment process.

The above rms value may represent either the rms residual associated with the observational grid, or the rms of the predictions computed in a similar manner (at the data generating phase, errorless predictions in a "prediction grid" are also generated, to be later compared with the predictions obtained from the adjustment). For the sake of economy in such evaluations, the rms values will be computed using only a part of the appropriate grid and not every point of the grid. However, even then two significant digits will be usually safeguarded.

One further approximation, geared toward simplifying the analysis, is utilized. It has been observed that the rms values associated with observation points and those associated with prediction points are usually quite similar. This statement holds with respect to both geoid undulations and gravity anomalies, and its validity is not substantially impared by the removal of constraints and/or by changes in the depth of point masses. An adjustment of gravity anomalies in which the point masses are in a much shallower configuration than usual could be an exception in this respect. However, since in most cases examined the two kinds of rms differ by less than 10% (usually in favor of the observation points), an average value will be taken and referred to as "rms fit".

Although in most practical cases dealing with a local representation of the gravity field we would be concerned with modeling of geoid undulations obtained via satellite altimetry, the algorithm for modeling of gravity anomalies will be analyzed in detail as well; it could serve in certain specialized cases or in view of some future applications. The P.M. parameters will be used not only for predictions of the same quantity as the one being adjusted (e.g., for predictions of geoid undulations based on an adjustment of geoid undulations), but also for predictions of a different quantity (e.g., for predictions of gravity anomalies based on an adjustment of geoid undulations). In this way, one circumvents the need for a covariance function or a cross-covariance function, both of which are employed in conjunction with some other local models (such as a model formulated in terms of node points) and may be obtained either from theoretical formulas by averaging over the whole globe, or from some empirical local formulas, etc. Such functions express the dependence of one quantity on the same quantity at a different point, or on a different quantity.

The observations will be generated so as to approach an ideal distribution as closely as practicable; their density will correspond to approximately two observations per point mass in each direction as already pointed out in the introductory section. Due to the characteristics of the S.H. model, the first (global) adjustment will yield essentially the same values of potential coefficients and thus of predicted values and adjusted observations, whether the (errorless) observed quantities are geoid undulations or gravity anomalies. The sigmas will be, of course, different since the a-priori sigma of 1m for geoid undulations is several-fold superior to 6 mgal used in conjunction with gravity anomalies.

Two basic adjustment groups will be analyzed. The first group, representing a global type of adjustment, will contain geoid undulation adjustments as well as gravity anomaly adjustments. The uniform grid of generated observations and the grid of point masses in the second adjustment will extend over the whole globe. Although the global distribution of point masses is only of theoretical value, the results will be a good indication of accuracy which could be achieved in more practical situations.

The second group, representing a global adjustment followed by a local adjustment, will complement the above investigations. The observations will be represented by a high density grid of geoid undulations in a limited area and by a low density global grid of gravity anomalies whose primary role is to make the first adjustment in terms of the S.H. potential coefficients at all possible. The point masses in the second adjustment will be distributed within the high density observational grid. The concept of local representation of the geoid surface favors a relatively high degree and order

S.H. model in the first adjustment, as opposed to holding the S.H. coefficients fixed or adjusting only a few very low degree and order coefficients. The main point of this argument is that a higher degree and order S.H. model would accommodate the local undulation data reasonably well, thus the resulting residuals would be much smaller than in the other cases and would have less of a systematic character. We could also say that such a model represents a better trend function.

A better trend function means that

a) the mass or volume constraints to be applied in a subsequent P.M. adjustment are more meaningful (if the S.H. "geoid" were not adjusted to fit the local data it could depart significantly, over the region of interest, from the observed geoid; the average bias would be enforced by the constraints -- even if only a reduced number of constraints were used -- and thus could not be removed by the point mass parameters), and

b) the smaller size of an average residual ensures that the spherical approximation embodied in the P.M. model becomes less important or completely insignificant.

These two points of view are relevant especially for the second group (in which a global S.H. adjustment is followed by a local P.M. adjustment) because a given S.H. model can accommodate a dense grid of observations extending over one region (and nowhere else) much better than it could accommodate the same observations in the presence of additional observations distributed over the rest of the globe.

The analysis concerned with the first group (which will be called "global case") includes both the original design and the dense design. The original design will eventually be eliminated because of the low accuracy

in the results. The results obtained from a simultaneous adjustment of the S.H. and P.M. parameters will not be listed since this approach would be of little practical value, due to a large number of parameters to be adjusted. As expected, a slightly better rms fit is obtained from such an adjustment than from two separate adjustments analyzed herein. In the second group (which will be called "local case"), only the analysis of the dense design will be included. The typical results for both groups follow. The rms fit for geoid undulations is denoted as rms(N), and that for gravity anomalies is denoted as rms(Δg); the depth of the point masses is referred to simply as "depth".

Global Case

Original Design, Geoid Undulations

a) When the depth increases, the rms(N) decreases but the correlations increase. When the depth becomes much greater than d, such as 2d, the correlations become exceedingly high; in such cases, systematic errors in observations located in one area could adversely affect the resulting geoid, etc., in large surrounding areas. On the other hand, as the depth decreases, the rms(N) increases rapidly. With regard to the predictions, the rms(Δg) follows a similar pattern. The best choice thus seems to be

$$\text{depth} = d. \tag{5.17}$$

As an illustration, results associated with three selected depths (½d, d, 2d) are listed below, where the errorless data correspond to a (12,12) S.H. model, the first adjustment to an (8,8) S.H. model, and the second adjustment to 184 point masses:

S.H.(8,8) ... rms(N) = 5.1m, rms(Δg) = 7.3 mgal;

P.M.(184) { ½d ... rms(N) = 2.4m, rms(Δg) = 6.3 mgal;
d ... rms(N) = 1.5m, rms(Δg) = 3.2 mgal;
2d ... rms(N) = 0.9m, rms(Δg) = 1.7 mgal.

b) The lowest rms(N) is obtained with no constraints. The volume constraints affect the rms(N) only slightly, while the mass constraints (and, of course, volume and mass constraints applied together) increase this value significantly. A similar conclusion holds also for the rms(Δg). Item c below will further illustrate this point.

c) The inclusion of higher order and degree S.H. coefficients is helpful in the sense that the residuals entering a P.M. adjustment have in general smaller magnitudes. The final rms(N) is then also proportionately lower. Some further results associated with the example of item a are the following (the number of P.M. parameters introduced in the second adjustment is 184 and their depth is d; the results, from left to right, pertain to no constraints, mass constraints, volume constraints, and mass constraints together with volume constraints):

(1) S.H.(8,8) ... rms(N) = 5.1m, rms(Δg) = 7.3mgal;

P.M.(184) ... rms(N) = 1.5m, 1.9m, 1.5m, 2.1m,

rms(Δg) = 3.2mgal, 3.6mgal, 3.2mgal, 3.7mgal;

(2) S.H.(10,10) ... rms(N) = 3.3m, rms(Δg) = 5.3mgal;

P.M.(184) ... rms(N) = 1.2m, 1.5m, 1.2m, 1.6m,

rms(Δg) = 2.5mgal, 2.6mgal, 2.5mgal, 2.7mgal.

In theory, the constraint blocks in (1), (2) should be different from each other, depending on the underlying S.H. model; however, in the two examples above, these blocks are larger than those stipulated, corresponding to a (6,6) model in both cases. It is to be noted that if the depth is 2d, the differences between the rms of the examples (1) and (2) are much smaller.

d) In this design, the P.M. parameters reduce the rms(N) obtained in a S.H. adjustment to about one third. Consequently, S.H. coefficients of an exceedingly high degree and order would have to be utilized before the results would even start approaching a decimeter accuracy with respect to a chosen resolution. Such a design would therefore present important practical disadvantages in actual data reductions.

Original Design, Gravity Anomalies

e) Item a applies also when gravity anomalies are being adjusted, but only with respect to the rms(Δg); in fact, the counterparts of such values listed in item a would now be slightly lower (6.3 mgal, 3.2 mgal, and 1.7 mgal would be replaced by 5.6 mgal, 2.9 mgal, and 1.5 mgal, respectively). The rms(N) will be treated separately in item i.

f) The lowest rms(Δg) is obtained with no constraints. The volume constraints affect the rms(Δg) only slightly while the mass constraints increase this value more significantly. Item g below will further illustrate this point.

g) Similar to item c, the inclusion of higher order and degree S.H. coefficients is helpful in reducing the rms(Δg). Furthermore, the values of rms(Δg) as presented below compare well with rms(Δg) of item c. In this respect we have

   (1) rms(Δg) = 2.9mgal, 3.1mgal, 2.9mgal, 3.5mgal;

   (2) rms(Δg) = 2.3mgal, 2.4mgal, 2.3mgal, 2.6mgal.

h) In the present design, the P.M. parameters reduce the rms(Δg) obtained in a S.H. adjustment to between one third and one half. Therefore this design does not appear suitable even if the gravity anomalies alone should be the values to be adjusted and predicted.

i) The situation with respect to predicted geoid undulations is further complicated by the fact that an adjustment without constraints would yield completely absurd results. In particular, it has been observed that

   (1) without constraints, the rms(N) would be extremely high (in some cases several hundred meters); although the rms(Δg) would be improving with increasing depth, the rms(N) would be rapidly deteriorating;

(2) the best results would be obtained in the presence of the volume constraints alone, in which case the values of rms(N) and rms(Δg) would be very similar to those obtained when adjusting geoid undulations subject to the same volume constraints;

(3) in conclusion, the present design would be unsuitable not only for an adjustment of geoid undulation but also for an adjustment of gravity anomalies, whether with regard to the predictions of gravity anomalies themselves (see item h), or for the predictions of geoid undulations (see (2) above in connection with item d).

Having discarded the original design from further considerations, we now turn to the dense design.

Dense Design, Geoid Undulations

j) Item a applies, together with the choice expressed in (5.17). As an illustration, results associated with two selected depths ($\frac{1}{2}$d,d) are listed below, where the errorless data correspond to a (5,5) S.H. model, the first adjustment to a (4,4) model, and the second adjustment to 128 point masses:

S.H.(4,4) ... rms(N) = 6.1m, rms(Δg) = 3.8mgal;

P.M.(128) { $\frac{1}{2}$d ... rms(N) = 0.18m, rms(Δg) = 0.36mgal,
d ... rms(N) = 0.01m, rms(Δg) = 0.02mgal.

As the depth decreases from d to $\frac{1}{2}$d, each rms increases about 20-fold. For a model containing higher degree and order S.H. coefficients, the increase in a similar situation would still be several-fold.

k) Either of the two kinds of constraints increases the rms(N) and rms(Δg) significantly.

(1) Returning to the example in item j for the depth d, we have

S.H.(4,4) ... rms(N) = 6.1m , rms(Δg) = 3.8mgal;

P.M.(128) ... rms(N) = 0.01m, 0.20m, 0.56m, 0.58m,

rms(Δg)= 0.02mgal, 0.21mgal, 0.23mgal, 0.27mgal.

(2) As another example, the errorless data now correspond to a (6,6) S.H. model, the first adjustment to a (4,4) S.H. model, and the second adjustment to 184 point masses at the depth d. This yields

S.H.(4,4) ... rms(N) = 8.3m, rms(Δg) = 5.6mgal;

P.M.(184) ... rms(N) = 0.01m , 0.12m, 1.0m, 1.0m,

rms(Δg)= 0.03mgal, 0.21mgal, 0.5mgal, 0.5mgal.

(3) The lowest rms(N), rms(Δg) are again obtained with no constraints present in the adjustment. However, the constraints least detrimental to these rms values are now the mass constraints.

(4) If the sizes of the constraint blocks are increased twice in each dimension, the detrimental effect of the constraints on the rms values is reduced to about 10% - 20% of the original effect.

l) A higher degree and order truncation in a S.H. model can again improve the rms(N), rms(Δg). Here comparison can be made with (2) of item k as follows:

S.H.(5,5) ... rms(N) = 5.1m , rms(Δg) = 3.9mgal;

P.M.(184) ... rms(N) = 0.01m, 0.09m, 0.8m, 0.8m,

rms(Δg)= 0.03mgal, 0.17mgal, 0.4mgal, 0.4mgal.

The version with no constraints gives almost perfect results in both cases being compared (0.008m versus 0.006m).

m) In this design, the desired geoidal detail can be expressed very well with no constraints and quite well, in terms of both the rms(N) and rms(Δg), with only the mass constraints present, especially if the size of the constraint blocks is increased (this leads to a smaller number of constraints). The degree and order of the S.H. model investigated has been admittedly low (only of degree and order 6,6 at the data generating level), mostly due to practical considerations. However, a more realistic case will be presented when dealing with the local case where the number of P.M. parameters can still be manageable even for a relatively high resolution of the geoidal features.

Dense Design, Gravity Anomalies

n) Item j applies also when gravity anomalies are being adjusted, but only with respect to the rms(Δg); in fact, the counterparts of the rms(Δg) values listed for $\frac{1}{2}d$,d in item j would now be slightly lower (0.36mgal, 0.02mgal would be replaced by 0.35mgal, 0.01mgal, respectively). The rms(N) will be treated separately in item r.

o) Item k applies, but only with respect to the rms(Δg). When Δg is being adjusted, the rms(Δg) are smaller than their counterparts from (1) and (2) of item k where N was the quantity to be adjusted. In this respect, we have

(1) rms(Δg) = 0.01mgal, 0.17mgal, 0.13mgal, 0.20mgal;

(2) rms(Δg) = 0.01mgal, 0.16mgal, 0.20mgal, 0.27mgal.

(3) The mass and the volume constraints exercise now about the same influence in increasing the rms(Δg).

(4) If the depth is d or larger than d and if the sizes of the constraint blocks are increased twice in each dimension, a several-fold improvement in the rms(Δg) can be expected (on the other hand, it has been observed that this rms does not improve for the depth of $\frac{1}{2}d$).

p) A higher degree and order truncation in a S.H. model can improve the rms(Δg). In the example below, in which a (5,5) instead of a (4,4) S.H. model is utilized in the first adjustment, a comparison with (2) of item o reveals an improvement with respect to the volume constraints alone or in combination with the mass constraints:

rms(Δg) = 0.01mgal, 0.16mgal, 0.15mgal, 0.20mgal.

These rms values are again lower than their counterparts in item l dealing with an adjustment of geoid undulations.

q) If we were interested strictly in predicting the Δg values, the option with no constraints would probably be the most natural choice. However, acceptable results could still be produced with mass or volume constraints, especially if the latter were applied in larger constraint blocks than stipulated.

r) The first part of item i applies also in the present situation. In particular,

(1) the statement in (1) of item i can be repeated, indicating that an unconstrained adjustment would be worthless insofar as N is concerned;

(2) the mass constraints are of little help, leaving the rms(N) too high (several meters) for any useful purpose; the volume constraints help in bringing the rms(N) down to approximately a one-meter level (this type of accuracy was obtained also when the same volume constraints were applied in the adjustment of geoid undulations);

(3) constraints applied in larger blocks are of little value;

(4) accordingly, even when assuming the application of volume constraints, the geoid determination in this category is of a relatively low accuracy (about one meter).

It appears that the use of point masses for the determination of a local geoid from local gravity anomalies lacks in accuracy. The treatment of mostly local gravity anomaly data would certainly lead to even worse results than those in the above adjustment of global data (with ideal characteristics) which already exhibits serious shortcomings in this respect. Therefore, when a local determination of the geoid surface from local data is contemplated, the attention should be focused on the adjustment of geoid undulations. In the remainder of this section we shall attempt to confirm the basic result of item m also in the limited area case.

We shall now describe an example set up to illustrate a fairly realistic local type of adjustment. The errorless data is generated by a set of S.H. potential coefficients complete through the degree and order (12,12). The first adjustment is performed in terms of a (6,6) S.H. model

and is followed by the second adjustment of 48 P.M. parameters. The global data making the first adjustment possible consists of mean gravity anomalies in $15^0 \times 15^0$ equilateral blocks (this corresponds to two observations in each direction per shortest half wavelength of about 30^0 as is expressed by the 6,6 model). The shortest half wavelength generated by the errorless (12, 12) model and expected to be recovered by the P.M. model is about 15^0; accordingly, the point masses are spaced at about 7.5^0 intervals (in an equilateral grid), covering a region of about $38^0 \times 45^0$. The local observations of geoid undulations are distributed in a $4^0 \times 4^0$ grid, about twice as dense in each dimension as the P.M. grid. The P.M. grid is well within the observational grid. According to (5.16), the constraint blocks should extend over about $60^0 \times 60^0$ (equilateral) regions, in agreement with $\bar{\theta}^0 = 180^0/\bar{n}$ where $\bar{n} = 6$. Our block containing all the point masses is substantially smaller, but since no overlapping blocks are used, the theoretical number of constraints is not substantially altered (we are in the presence of only one mass constraint and one volume constraint).

In the first adjustment, all the (global) Δg values and all the (local) N values are considered. The local geoid undulations are accommodated much better than they would be in the presence of other such observations outside the area of interest. Thus, the rms(N) after the first, (6,6) S.H. adjustment is relatively small (it is equal to 2.4m). Due to the fact that one area contains all the observed geoid undulations, the predictions outside this area are completely meaningless, whether they are obtained in the first or in the second adjustment. As could be expected, the second adjustment yields the best rms(N) values for points within the area containing the point masses. These will also be the representative values

presented; the rms(Δg) will have a limited reliability since only a small sample of predictions will be used in its computation. Since only the values at prediction points will be used in these computations, the rms values will be on the conservative side (it is believed that a reduction of 10% in these values would be achieved if both the prediction and the observation points were considered).

The savings in terms of the number of parameters may be of interest to practitioners. If a (12,12) S.H. model were to be adjusted to express the desired geoidal detail, the total number of parameters would be 169. However, the first adjustment in our procedure involves only 49 parameters (S.H. coefficients) and the second adjustment involves another 48 parameters (P.M. magnitudes), for a total of 97 parameters; however, since each set of parameters is adjusted separately, further reduction in the computing burden is achieved.

We shall now express a few additional items with regard to the second, P.M. adjustment in our example.

Local Case

Dense Design, Geoid Undulations

s) On the whole,item a applies again, together with the choice expressed in (5.17). However, the rms(Δg) now does not improve with the increasing depth. With regard to the rms(N), the following results have been obtained:

$\frac{1}{2}d$ ... rms(N) = 0.37m ,

d ... rms(N) = 0.32m ,

$(5/4)d$ ... rms(N) = 0.26m ,

$2d$ ... rms(N) = 0.20m .

t) The mass constraint (encompassing the whole P.M. block) and/or the volume constraint do not seem to exercise much influence on the rms(N). On the other hand, the mass constraint could have a beneficial effect on the rms(Δg), however, this has been confirmed only by a limited sample. This finding is in partial agreement with the statement in (3) of item k. The results obtained with the depth equal to d are:

rms(N) = 0.32m, 0.33m, 0.32m, 0.35m ,

rms(Δg)= 0.4 mgal, 0.2 mgal, 0.3 mgal, 0.2 mgal.

u) Accordingly, the overall best results in this category are those obtained either in the absence of constraints, or with the mass constraint alone, the depth being equal to d as in the previously investigated cases. In summary, the present investigation has yielded the following outcome:

rms(N) = 0.3m (a conservative estimate),

rms(Δg) = 0.2 - 0.4mgal (only a small sample involved).

5.5 Conclusion

This study had originally started from the notion that if the globe were covered with an ideal observational grid and if the point mass parameters had some of the desirable properties of the spherical harmonics, the total number of parameters should be invariant, in the global case, of the chosen degree and order of the spherical harmonic part of the complete model. This number should be then also equal to the number of parameters in such a spherical harmonic model which, alone, would produce about the same final resolution. Similar considerations suggest that on the average, one constraint per $\Delta\bar{S}$ equal area block would be desirable in the regions containing point masses; the sides of $\Delta\bar{S}$ would correspond to a half wavelength of the underlying spherical harmonic resolution. The actual arrangement of the constraints would be such that a constraint block ΔA would be four times as large as $\Delta\bar{S}$, but each ΔA block would involve two constraints (one called mass constraint and the other, volume constraint) and the overlapping constraint blocks would about double the number of the original ΔA blocks (see also equation 5.16 and the text below it). The horizontal separation of the point masses was considered to correspond to s, the shortest half wavelength of the geoidal features which should be taken into account by the adjustment. The depth (d) of these point masses beneath the surface of a reference ellipsoid (or a reference sphere in the usual case where the spherical approximation is utilized in the formulas) was stipulated as equal to 0.8s.

In reality, however, the point mass parameters are not mutually independent even if both the observations and the point masses are distributed in an idealized grid, whether on the whole globe or a portion thereof, and even if the point masses are much shallower than originally stipulated. This points to the need to increase the number of these parameters. The attempts to achieve a reasonable accuracy in reproducing a "true" geoid have indicated that the grid of point masses should be about twice as dense along each dimension as the one originally stipulated. This has lead to the arrangement called here the "dense design". The analysis of the previous section has indicated that this design is potentially capable of resulting in a decimeter accuracy in the geoid determination.

Although the "original design" had to be abandoned because of the low accuracy in the final results, the depth of point masses such that $d = 0.8s$ has been shown to be, overall, the most advantageous. It has also been suggested in the previous section that the first adjustment (in terms of a spherical harmonic model) should include as high degree and order potential coefficients as practicable. The point masses in the subsequent adjustment are to be employed only in areas of a relatively high observational density, thus assuring that the number of parameters (point mass magnitudes) may be kept reasonably small.

The analysis of the previous section has resulted in a few suggestions regarding the accuracy that could potentially be achieved in the above two adjustment steps. If the observations consisted of mean gravity anomalies (evenly distributed over the whole globe, having about the same quality, etc.), reasonably accurate predictions of gravity anomalies could be made with

no constraints, or with mass and/or volume constraints applied in larger constraint blocks than stipulated. On the other hand, geoid undulations could be predicted with any success,in this gravity anomaly adjustment, only with the use of volume constraints, and even then the accuracy would be relatively low (on the order of one meter).

If the globe were covered with well distributed observations consisting of geoid undulations, an adjustment of these quantities in the two steps described would yield good results (approaching perhaps a decimeter accuracy in the geoid determination), both for predictions of geoid undulations and gravity anomalies. The best accuracy would be achieved with no constraint at all, or with only the mass constraints, especially if larger constraint blocks than those stipulated were utilized.

The most practical arrangement is, of course, that of mean gravity anomalies available over the whole globe (with some exceptions), and of densely distributed geoid undulations available over a limited area. The latter are usually the result of an adjustment of satellite altimeter data. In such a situation, good results in terms of both geoid undulations and gravity anomalies could be obtained either without any constraints at all, or perhaps with one mass constraint covering the region of interest; this constraint would ensure that the sum of all point mass magnitudes present in the second adjustment is zero. It appears that a decimeter accuracy for a desired resolution can be approached if the point mass adjustment is based on a fairly high degree and order spherical harmonic model (e.g., 12,12 or 14,14 model) and if the geoidal detail to be represented by such an adjustment is contained well within the point mass region which, in turn, should be well within the region of high density observations.

6. CONCLUSIONS

In several respects, this report has been a natural continuation of the effort described in [Blaha, 1977]. The research contained herein has been based on the concept developed in recent years, in which the short arc model of satellite altimetry has been combined with the gravity anomaly model in the determination of the geoidal parameters (spherical harmonic potential coefficients); the same simultaneous adjustment has also yielded the revised values of weakly constrained state vector parameters defining independent, interlocking short arcs.

The main thrust of the present effort has been directed at improving the knowledge of the earth's gravity field and of its fundamental surface, the geoid. This improvement can be realized along two lines, by upgrading the accuracy of the existing algorithms and by providing for a more detailed representation of the local geoid in areas where a dense configuration of altimeter observations is available. The latter task can be accomplished by supplementing the geoidal model based on the potential coefficients with a new kind of parameters represented by point masses at chosen locations. Both lines of research will be briefly summarized in the following paragraphs.

In the process of adjusting satellite altimeter data, the radial distance (length of a position vector) from the geocenter to the geoid as defined by the spherical harmonic potential coefficients is needed. The

geocentric latitude and longitude associated with this distance are obtained when forming the observation equations. In the past, the radial distance has been computed to a desired accuracy in an iterative process. Even if a crude initial value is adopted, a sub-meter accuracy is achieved on the second iteration, while the third iteration yields a sub-millimeter accuracy. If the best possible initial value is taken, such as the radial distance to the mean earth ellipsoid, the iterative process may be accelerated by one iteration. But even then two iterations would be needed in most cases. However, an algorithm derived and described in Chapter 2 yields a sub-centimeter accuracy already from the first iteration. It results in important computer savings, considering that in real data reductions of satellite altimetry the radial distance needs to be computed at thousands of locations. In most practical cases where an iterative solution would not be attempted for economical reasons, this new algorithm symbolizes a significant increase in accuracy.

A key element in any attempt to achieve a decimeter accuracy in geoid representation via satellite altimetry is the necessity either to obtain the ephemeris of comparable accuracy, or to circumvent this requirement by adjusting the ephemeris in some way, together with the geoid. The first possibility requires extensive satellite tracking and involves an enormous number of adjustable parameters in the long arc approach. The second approach, adopted in this study, allows for a piece-wise treatment of short orbital arcs considered mutually independent, in which slight adjustments of the state vector parameters can compensate for an inherent modeling error. The main question which had to be answered when pondering

the possibility of using the short arc adjustment model in altimetry reductions involving SEASAT or a similar observational system is whether or not this method is inherently capable of representing the geoid to the desired accuracy. An analysis presented in Chapter 3 has provided at least a partial answer to this question by pointing out the necessary conditions for accomplishing this goal; under favorable circumstances, these conditions could prove to be also sufficient. They consist almost exclusively in specifying the maximum length of a sub-arc compatible with the given accuracy requirements.

The analysis presented in the appendix of this report contains a proof of a basic equivalence, at the level of observation equations, between the gravity anomaly model and the gravity model, both considered at the same geoidal point and expanded in terms of spherical harmonics. However, a close scrutiny reveals the following fine distinction. The gravity anomaly model contains two approximations, called the spherical and the direction approximations. The gravity model, on the other hand, contains only the direction approximation which is the less important of the two. The accuracy of the gravity anomaly model can thus be increased upon its transformation to the corresponding gravity model. Since either of the above approximations introduces a very small error (the larger, spherical error can hardly reach 0.2mgal), their analysis appears to be mostly of theoretical interest. However, the refinement of the gravity anomaly model could become desirable in the future, in view of higher data density and more accurate measurements. The most important outcome symbolizing this refinement is reviewed and discussed in Chapter 4.

The purpose of introducing point mass parameters into an adjustment of gravity and satellite altimetry is to add fine structure to a geopotential model based on the (adjustable) spherical harmonic coefficients. The point mass adjustment, described in Chapter 5, has been conceived to follow an adjustment of satellite altimetry and gravity anomalies performed in terms of spherical harmonics. The point masses are to be employed only in areas of a relatively high observational density, thus assuring that the number of parameters (point mass magnitudes) may be kept reasonably small.

The original notion of the point mass model indicates the presence of two kinds of constraints called mass constraints and volume constraints. In the past this model was idealized, in the sense that the point mass parameters were assumed, under certain conditions, to be mutually independent, similar to the case with the spherical harmonic coefficients. The idealized model lead further to the stipulation that the horizontal separation of the point masses should be approximately equal to the shortest half wavelength (s) of the desired resolution.

In reality, however, the above independence property does not hold even if both the observations and the point masses are distributed in an idealized grid, whether on the whole globe or a portion thereof, and even if the point masses are much shallower than originally stipulated. This points to the need to increase the number of these parameters. The attempts to achieve a reasonable accuracy in reproducing a "true" geoid have indicated that the grid of point masses should be about twice as dense along each dimension as the one originally stipulated. The most advantageous depth of point masses has been found as 0.8s. In the practical case of fitting the

geoid undulations (as derived from satellite altimetry) using the point mass model, no constraints appear to be necessary. Sometimes one mass constraint, which assures that the sum of all the point mass magnitudes introduced in such an adjustment is zero, may be beneficial for the predictions of gravity anomalies. On the other hand, it appears that the predictions of geoid undulations obtained in the adjustment of gravity anomalies via point masses would exhibit insufficient accuracy in most cases. A more detailed accuracy analysis is contained in Chapter 5.

APPENDIX

REFINEMENT OF THE GRAVITY ANOMALY MODEL

A.1 Statement of the Problem and Objectives

In the first part of this analysis, our attention will be focussed on the derivation of the gravity anomaly model and the gravity model, both considered at the same geoidal point and expanded in terms of spherical harmonics. In the latter part, we shall form the observation equations for the two models and show that for most practical purposes they are identical. They should thus lead to nearly identical results not only with regard to the final values of the parameters (spherical harmonic potential coefficients), but also with regard to predictions and the variance-covariance propagation. If for some reason more accuracy is required in the customarily used gravity anomaly adjustment model, one way of accomplishing such a goal will be indicated.

The mathematical models will be expressed in terms of the conventional spherical harmonic potential coefficients C_{nm} and S_{nm} (C's and S's) among which will be included the "central term" r_o, at least in theoretical considerations. The geoidal point (P), at which the gravity is denoted by g and the gravity anomaly by Δg, has the following spherical coordinates:

$$P \ldots \bar{\phi}, \lambda, r, \tag{A.1a}$$

where, symbolically,

$$r = r \text{ (potential coefficients; } \bar{\phi},\lambda) \qquad \text{(A.1b)}$$

is the radial distance from the coordinate origin to the $\bar{\phi},\lambda$ location on the geoid as determined by the C's and S's, and where

$$\bar{\phi} = \text{arc tg } [(1 - e^2) \text{ tg } \phi],$$

ϕ being the geodetic (or ellipsoidal) latitude and e being the eccentricity of the reference ellipsoid. Of the three spherical coordinates, only r is a function of the potential coefficients. It may be of interest that in terms of C's and S's, the gravity model for g' associated with $\bar{\phi},\lambda$, $r' = r+h$ would be equivalent to the model for g associated with $\bar{\phi},\lambda,r$ provided g is obtained from g' through the free-air reduction.

The potential at P is given by an infinite series which, in practice, is truncated at some suitable $n = N$. This constitutes a reason why less concern has been attached recently to the convergence problems associated with such series (see e.g. [Hotine,1969], page 173; [Needham,1970], pages 46 and 47; or [Rapp,1972], page 20). Most of the time, the formulas in geodetic literature imply that

$$C_{10} = C_{11} = S_{11} = 0 \qquad \text{(A.2a)}$$

and the Cartesian origin coincides with the earth's center of mass. This means that there are no first-degree terms in the spherical harmonic expansion of the potential. Furthermore, the usual form of the contribution of the earth's rotation in the formula giving the potential is consistent with

the condition

$$C_{21} = S_{21} = 0, \tag{A.2b}$$

specifying that the Cartesian z-axis is parallel to the earth's axis of rotation. In fact, combined with (A.2a), this condition implies that the two axes coincide. The equations (A.2) represent the usual assumptions related to the definition of the coordinate system and the harmonics corresponding to the above five coefficients are sometimes called "forbidden harmonics" (see e.g. the discussion on page 33 of [Blaha,1975]).

When comparing the expansions for the actual and the normal potentials, we shall utilize the same symbols kM (the gravitational constant times the earth's mass) and ω (the earth's rotation rate) throughout. For it is assumed that the reference ellipsoid has the same mass and the same angular velocity as the actual earth. We also assume, at least initially, that the reference ellipsoid, on which $U = U_0$ (U is the normal potential), and the geoid, on which $W = W_0$ (W is the actual potential), have the same numerical value of the potential, namely

$$W_0 = U_0 \ . \tag{A.3}$$

Since an equipotential ellipsoid of revolution and its gravity field are completely determined by four constants, there is only one ellipsoid, called the mean earth ellipsoid, which shares with the actual earth the four parameters kM, ω, W_0, and $J_2 = -C_{20}$. The ellipsoid thus defined is the best global approximation of the earth by an ellipsoid (see [Moritz,1975], pages 136,137). In particular, from [Heiskanen and Moritz, 1967], pages 109,110,

it follows that the sum of the squares of the deviations N of the geoid from the mean earth ellipsoid is a minimum, and in an absolute position (when its center coincides with the earth's center of mass), the mean earth ellipsoid is particularly well suited for dynamical astronomy. If we adopt J_2 of the earth in the definition of the reference ellipsoid it becomes clear, due to the previous three assumptions, that the normal gravity field is defined in terms of the mean earth ellipsoid. As a consequence of the equations (A.2), this ellipsoid is in an absolute position and its z-axis coincides with the earth's axis of rotation. The definition of the ellipsoid enters into the values of gravity anomalies through the values of the constants a, b, γ_e, γ_p (the semi-major and semi-minor axes, the equatorial and the polar gravity), needed in order to compute γ (the normal gravity) at points on its surface according to the formula:

$$\gamma = (a\gamma_e \cos^2\phi + b\gamma_p \sin^2\phi)/(a^2 \cos^2\phi + b^2 \sin^2\phi)^{\frac{1}{2}} .$$

In this rigorous formula (Somigliana's formula) the symbol ϕ represents the geodetic latitude.

From the adjustment point of view, the parameters kM and ω as given for the reference ellipsoid (here represented by the mean earth ellipsoid) will be held fixed, i.e., the earth's mass and rotation rate will not be subject to adjustment. On the other hand, we may or may not decide to adjust W_o or C_{20}. Thus, ideally, we could arrive at a "better mean earth ellipsoid". In practice, however, we can only arrive at an ellipsoid which may suite somewhat better a particular set of data.

The parameter W_o deserves further attention. We replace it by another parameter (r_o) which represents the radius of a ficticious spherical earth having a homogeneous mass distribution and having the same potential as the geoid; in particular,

$$r_o = kM/W_o \ . \qquad \text{(A.4a)}$$

It is now this parameter that may or may not be subject to adjustment. The corresponding fixed quantity related to the reference ellipsoid is

$$r_o^* = kM/U_o \ ; \qquad \text{(A.4b)}$$

further, we introduce the following "associated parameter":

$$\Delta r_o = r_o - r_o^* \ . \qquad \text{(A.4c)}$$

From this point on, we shall consider r_o or Δr_o subject to adjustment. We can then define r_o^* to have the role of a starting (approximate) value of the parameter r_o; U_o would similarly be considered as a starting value for W_o. Should r_o or Δr_o be held fixed, the parameters $r_o = r_o^*$ or $\Delta r_o = 0$ would be assigned an extremely large weight or, equivalently, the corresponding rows and columns would simply be deleted from the normal equation matrices.

The series expressing the normal potential contains only the $J_2, J_4, \ldots$ coefficients, or equivalently (except for the sign), only the $C_{20}^*, C_{40}^*, \ldots$ coefficients, where the fixed C's in the normal field have been identified by "*". In practice, the coefficients beyond C_{40}^* are considered negligible (see e.g. [Rapp,1970], page 2). We shall in fact show that of C_{60}^* is replaced by zero, the largest possible error in the gravity anomaly

model is 0.03 mgal; with regard to the spherical harmonic expansion of normal gravity in the radial direction, it is 0.04 mgal (the contribution of C^*_{80} is two orders of magnitude smaller, etc.).

In analogy to our treatment of r_o, we find it convenient to accept C^*_{20} as a starting value for C_{20}; we also use the following notation for another associated parameter:

$$\Delta C_{20} = C_{20} - C^*_{20} \tag{A.5}$$

We note that if either r_o or C_{20} were completely free to adjust, any starting values sufficiently close to the final values would theoretically be acceptable, not only the ones (r^*_o, C^*_{20}) referring to the mean earth ellipsoid. On the other hand, should either of the parameters r_o or C_{20} be weighted at the values r^*_o or C^*_{20}, these would also be the most logical starting values creating no need for a special correction term on the right-hand side of the normal equations (we are concerned with one adjustment cycle only). On the whole, the use of the starting values r^*_o and C^*_{20} is desirable in any event. Whether r_o and C_{20} are completely free to adjust or not, such values are advantageous simply because they are usually close to the final values. In fact, the most recent parameters of the mean earth ellipsoid represent the best available information related to the actual earth.

Since only four parameters of the mean earth ellipsoid are theoretically equal to their counterparts related to the actual earth, C^*_{40} is not close in value to the actual C_{40} coefficient. We again denote an associated parameter as follows:

$$\Delta C_{40} = C_{40} - C^*_{40} \, . \tag{A.6}$$

Besides the cases (A.5), (A.6), all the other associated parameters which will be used in the gravity anomaly mathematical model are identical to the C's and S's themselves (C^*_{60}, etc., have been neglected), namely

$$\Delta C_{nm} = C_{nm}, \qquad (n,m) \neq (2,0),\ (4,0), \tag{A.7a}$$

$$\Delta S_{nm} = S_{nm}. \tag{A.7b}$$

The starting values of all the parameters ("regular" or "associated") are denoted by the superscript "0". Thus we have

$$(\Delta r_o)^0 = (r_o)^0 - r_o^*, \tag{A.8a}$$

$$\Delta C^0_{20} = C^0_{20} - C^*_{20}, \tag{A.8b}$$

$$\Delta C^0_{40} = C^0_{40} - C^*_{40}, \tag{A.8c}$$

$$\Delta C^0_{nm} = C^0_{nm}, \quad (n,m) \neq (2,0),(4,0), \tag{A.8d}$$

$$\Delta S^0_{nm} = S^0_{nm}. \tag{A.8e}$$

In accordance with the suggested procedure, most of the time we use

$$(\Delta r_o)^0 = 0, \tag{A.8a'}$$

$$\Delta C^0_{20} = 0. \tag{A.8b'}$$

When comparing the gravity anomaly adjustment model with the gravity adjustment model, we shall take advantage of the stipulation (A.8a').

A.2 Basic Formulas and Notations

The potential at point P having the spherical coordinates $(\bar{\phi},\lambda,r)$ is given by the formula

$$W = (kM/r)\left[1 + \sum_{n=2}^{N} (a/r)^n \sum_{m=0}^{n} (C_{nm}\cos m\lambda + S_{nm}\sin m\lambda)\, P_{nm}(\sin\bar{\phi})\right]$$

$$+ \tfrac{1}{2}\omega^2 r^2 \cos^2\bar{\phi}, \qquad (A.9)$$

where $P_{nm}(\sin\bar{\phi})$ represents the associated Legendre functions. Differentiating (A.9) in the radial direction and changing the sign, we obtain the radial component of gravity at P as follows:

$$g_r = -\partial W/\partial r,$$

$$g_r = (kM/r^2)\left[1 + \sum_{n=2}^{N} (n+1)(a/r)^n \sum_{m=0}^{n} (C_{nm}\cos m\lambda + S_{nm}\sin m\lambda)\right.$$

$$\left. \cdot P_{nm}(\sin\bar{\phi})\right] - \omega^2 r\cos^2\bar{\phi}. \qquad (A.10)$$

In considering the normal field (C^*_{60},C^*_{80}, etc., are neglected), we have for the same point P:

$$U = (kM/r)\left[1 + (a/r)^2 C^*_{20} P_2(\sin\bar{\phi}) + (a/r)^4 C^*_{40} P_4(\sin\bar{\phi})\right]$$

$$+ \tfrac{1}{2}\omega^2 r^2 \cos^2\bar{\phi}, \qquad (A.11)$$

$$\gamma_r = -\partial U/\partial r \ ,$$

$$\gamma_r = (kM/r^2)\,[1 + 3(a/r)^2\, C^*_{20}\, P_2(\sin\bar{\phi}) + 5(a/r)^4\, C^*_{40}\, P_4\,(\sin\bar{\phi})]$$

$$- \omega^2 r \cos^2\bar{\phi} \ . \qquad \text{(A.12)}$$

The greatest possible error in this expression caused by neglecting C^*_{60} is 0.04 mgal, in the evaluation of which we have used ± 1 for $P_6(\sin\bar{\phi})$, 980 mgal for kM/r^2, 1 for a/r, and -0.000 000 0061 for C^*_{60} (as given in GRS 1967). At point P, the disturbing potential (T) is defined as

$$T = W - U$$

which, in view of (A.9) and (A.11), yields

$$T = (kM/r) \sum_{n=2}^{N} (a/r)^n \sum_{m=0}^{n} (\Delta C_{nm} \cos m\lambda + \Delta S_{nm} \sin m\lambda)\, P_{nm}\,(\sin\bar{\phi}). \qquad \text{(A.13)}$$

Since P is considered to be on the geoid, W in (A.9) is replaced by

$$W_o = kM/r_o$$

according to (A.4a). Upon multiplying both sides of the new equation by $rr_o/(kM)$, we obtain

$$r = r_o \left[1 + \sum_{n=2}^{N} (a/r)^n \sum_{m=0}^{n} (C_{nm} \cos m\lambda + S_{nm} \sin m\lambda)\, P_{nm}\,(\sin\bar{\phi}) \right.$$

$$\left. + \tfrac{1}{2}\, \omega^2 r^3 \cos^2\bar{\phi}/(kM)\right]. \qquad \text{(A.14)}$$

The point on the reference ellipsoid corresponding to P is denoted as $\tilde{P}$; its geocentric latitude is taken, for our purpose, as $\bar{\phi}$ so that it is associated with the triplet $(\bar{\phi},\bar{\lambda},\tilde{r})$. In analogy to the above, the normal potential at $\tilde{P}$ is given by (A.11) where U is replaced by

$$U_o = kM/r_o^* \tag{A.15}$$

according to (A.4b), and r is replaced by $\tilde{r}$; the multiplication by $\tilde{r}\, r_o^*/(kM)$ yields

$$\tilde{r} = r_o^* \left[1 + (a/\tilde{r})^2 C_{20}^* P_2(\sin\bar{\phi}) + (a/\tilde{r})^4 C_{40}^* P_4(\sin\bar{\phi}) + \tfrac{1}{2}\omega^2\tilde{r}^3 \cos^2\bar{\phi}/(kM)\right]. \tag{A.16}$$

Some of the previous equations will be rewritten in a shorter form, upon using the notations introduced as follows:

$$S(n) = \sum_{m=0}^{n} (C_{nm} \cos m\lambda + S_{nm} \sin m\lambda)\, P_{nm}(\sin\bar{\phi}), \tag{A.17a}$$

$$\Delta S(n) = \sum_{m=0}^{n} (\Delta C_{nm} \cos m\lambda + \Delta S_{nm} \sin m\lambda)\, P_{nm}(\sin\bar{\phi}); \tag{A.17b}$$

$$P_1 = \sum_{n=2}^{N} (a/r)^n S(n), \tag{A.18a}$$

$$P_2 = \sum_{n=2}^{N} n(a/r)^n S(n), \tag{A.18b}$$

$$P_3 = \sum_{n=2}^{N} (n+1)(a/r)^n S(n), \tag{A.18c}$$

$$P_4 = \sum_{n=2}^{N} (n+1)(n+2)(a/r)^n \, S(n), \tag{A.18d}$$

$$P_5 = \sum_{n=2}^{N} (n-1)(a/r)^n \, S(n), \tag{A.18e}$$

$$P_6 = \sum_{n=2}^{N} (n+2)(n-1)(a/r)^n \, S(n); \tag{A.18f}$$

$$\Delta P_5 = \sum_{n=2}^{N} (n-1)(a/r)^n \, \Delta S(n) \equiv P_5 - (a/r)^2 \, C^*_{20} \, P_2(\sin\bar{\phi}) - 3(a/r)^4 \, C^*_{40} \, P_4(\sin\bar{\phi}), \tag{A.19a}$$

$$\Delta P_6 = \sum_{n=2}^{N} (n+2)(n-1)(a/r)^n \, \Delta S(n). \tag{A.19b}$$

We also abbreviate

$$C \equiv C(r) = kM/r^2, \tag{A.20a}$$

$$D \equiv D(r) = \omega^2 r^3 \cos^2\bar{\phi}/(kM) \equiv (\omega^2 r/C)\cos^2\bar{\phi}. \tag{A.20b}$$

No new notations will be introduced should the equations (A.17) through (A.20) represent numerical values computed with the starting values (p^0) of the parameters. Thus, the symbols $(P_1)^0$ or D^0, etc., will not be employed, but it will be clear from the context when the numerical values p^0 are involved. We shall now present some of the extreme values obtained with a set of "synthetic potential coefficients" used also in [Blaha,1977], complete through the degree and order (20,20). Since the relatively low degree and order coefficients are employed and since $a/r = 1$ is assumed, such values will be used only to estimate the errors committed by neglecting

certain terms when relating the gravity anomaly model to the gravity model. In the first three expressions below we indicate, besides the extreme values, also the corresponding location (here the south pole), and next to it we present the largest numerical values obtained in each case for a point on the equator:

$$P_1 \ldots \quad -0.0011 \ (\overline{\phi} = -90^0), \quad +0.00055 \ (\overline{\phi} = 0), \tag{A.21a}$$

$$P_2 \ldots \quad -0.0022 \ (\overline{\phi} = -90^0), \quad +0.0011 \ (\overline{\phi} = 0), \tag{A.21b}$$

$$P_4 \ldots \quad -0.0130 \ (\overline{\phi} = -90^0), \quad +0.0070 \ (\overline{\phi} = 0). \tag{A.21c}$$

Further we have

$$\Delta P_5 \ldots -0.000\,0042, \tag{A.22a}$$

$$\Delta P_6 \ldots -0.00057 \ . \tag{A.22b}$$

We shall also use

$$C \simeq 980 \text{ gal} \ , \tag{A.23a}$$

$$D \simeq 0.0034 \cos^2\overline{\phi} \ . \tag{A.23b}$$

We are now in the position to restate the equations (A.9), (A.10), (A.13), and (A.14) in a more convenient form, using the abbreviated expressions (A.17), (A.20); in the case of g_r and r, we further use (A.18c) and (A.18a), respectively. We thus have

$$W = (kM/r)\left[1 + \sum_{n=2}^{N} (a/r)^n \, S(n)\right] + \tfrac{1}{2}\omega^2 r^2 \cos^2\overline{\phi} \ , \tag{A.9'}$$

$$g_r = C\left[1 + \sum_{n=2}^{N} (n+1)(a/r)^n \, S(n)\right] - \omega^2 r \cos^2\overline{\phi} \equiv C\,(1 + P_3 - D) \tag{A.10'}$$

$$T = (kM/r) \sum_{n=2}^{N} (a/r)^n \Delta S(n) , \qquad (A.13')$$

$$r = r_o \left[1 + \sum_{n=2}^{N} (a/r)^n S(n) + \tfrac{1}{2}\omega^2 r^3 \cos^2\bar{\phi}/(kM)\right]$$

$$\equiv r_o(1 + P_1 + \tfrac{1}{2}D). \qquad (A.14')$$

In analogy to Δr_o defined in (A.4c), we may consider the following equivalent associated parameter:

$$\Delta W_o = W_o - U_o .$$

In fact, Δr_o will be introduced into an adjustment through ΔW_o. The relation between these two quantities is deduced from (A.15) as

$$\Delta W_o = -(kM/r_o^{*2})\, \Delta r_o . \qquad (A.24)$$

We shall finally derive certain differential relations which will be useful throughout. From (A.4c), we have

$$d\Delta r_o = dr_o , \qquad (A.25a)$$

and from (A.5), (A.6), (A.7) we deduce for any n,m:

$$d\Delta C_{nm} = dC_{nm} , \qquad (A.25b)$$

$$d\Delta S_{nm} = dS_{nm} . \qquad (A.25c)$$

The adjusted parameters (p^a) are obtained as

$$p^a = p^o + dp . \qquad (A.26)$$

In the case of associated parameters, one has

$$\Delta p^a = \Delta p^o + d\Delta p ,$$

which corresponds to (A.26) upon subtracting p* from both sides of the equation, knowing from (A.25) that $d\Delta p = dp$. The equations (A.25) at once yield

$$d\Delta S(n) = dS(n) \equiv \sum_{m=0}^{n} (dC_{nm} \cos m\lambda + dS_{nm} \sin m\lambda)\, P_{nm}(\sin\bar{\phi}). \tag{A.27}$$

Upon differentiating (A.14), we find

$$dr = (r/r_o)\, dr_o + (r_o/r) \left[\sum_{n=2}^{N} (-n)(a/r)^n S(n) + (3/2)D \right] dr + r_o \sum_{n=2}^{N} (a/r)^n\, dS(n) .$$

If the terms containing dr are grouped together and (A.18b) is recalled, we obtain

$$dr = \left[(r/r_o)\, dr_o + r_o \sum_{n=2}^{N} (a/r)^n\, dS(n)\right] / \left[1 + (r_o/r)(P_2 - 3D/2)\right] \tag{A.28}$$

This equation will be the backbone in showing the close similarity between the gravity anomaly model and the gravity model.

A.3 Gravity Anomaly Mathematical Model

When deriving a formula suitable for the adjustment of gravity anomalies, we shall rely heavily on the reference [Blaha,1977], henceforth abbreviated as [B]. Formulas in this reference are based on the implicit assumption that $\Delta W_o \equiv 0$. In our present approach, the main deviation from such a stipulation is that although numerically, as the starting value, we usually accept

$$(\Delta W_o)^0 = 0 \ ,$$

ΔW_o is in general treated as an adjustable parameter, i.e., $\Delta W_o \neq 0$. Keeping this in mind, we can go over the derivations in [B] with only the few changes listed below.

First, instead of $W_o = U_o$ (appearing on page 74 of [B]) we would now have

$$W_o = U_o + \Delta W_o \ .$$

The symbol T in the "specific Brun's formula" and in the second term of the exact boundary condition would accordingly be replaced by $(T - \Delta W_o)$. In going past the spherical approximation (unchanged), T in the second term of the approximate boundary condition would again be replaced by $(T - \Delta W_o)$, so that this condition (see the equation 5.12 in [B]) would now read

$$g_r \simeq -\partial T/\partial r - (2/r)(T - \Delta W_o) \ . \qquad (A.29)$$

As before (see the equation 5.11 in [B]), the definition of Δg_r may be written as

$$\Delta g_r = g_r - \gamma_r \, . \tag{A.30}$$

If we consider our formula (A.13) for T at a geoidal point P, upon performing $-\partial T/\partial r - (2/r)T$ we recover the right-hand side of (5.1) in [B], truncated at n = N, which gives a more usual expansion for Δg_r without the term Δr_o. We now have only to add the term

$$(2/r)\Delta W_o = (kM/r^2)(-2r/r_o^{*2})\Delta r_o \, ,$$

where (A.24) has been included. The "full" expansion then reads

$$\Delta g_r \simeq (kM/r^2) \left[-2(r/r_o^{*2})\Delta r_o + \sum_{n=2}^{N} (n-1)(a/r)^n \sum_{m=0}^{n} (\Delta C_{nm} \cos m\lambda + \Delta S_{nm} \sin m\lambda) P_{nm} (\sin\bar{\phi})\right]. \tag{A.31}$$

Taking advantage of our abbreviations, we write it as

$$\begin{aligned} \Delta g_r &\simeq C\left[-2(r/r_o^{*2})\Delta r_o + \sum_{n=2}^{N} (n-1)(a/r)^n \, \Delta S(n)\right] \\ &\equiv C\left[-2(r/r_o^{*2})\Delta r_o + \Delta P_5\right]. \end{aligned} \tag{A.31'}$$

We notice that the greatest possible error caused by neglecting C_{60}^* in this expression amounts to no more than 0.03 mgal. In fact, using values similar to those presented following the equation (A.12), we confirm that

$$-(980 \text{ gal})\, 5\, C_{60}^* (\pm 1) = \pm 0.03 \text{ mgal}.$$

We shall next examine the change in Δg_r due to the change (dp) in the parameters from their starting values (p^0). We propose to do so by first considering r fixed and then adding a correction (c) due to r being a function of the parameters. During this process we shall take advantage of (A.25). In the first step (in which Δg_r is linear in the parameters) we obtain

$$d\Delta g_r \simeq C\left[-2(r/r_o^{*2})\, dr_o + \sum_{n=2}^{N} (n-1)(a/r)^n\, dS(n)\right] . \qquad (A.32)$$

If, when differentiating (A.31') with respect to r in the second step, we adopt zero for the starting value of Δr_o according to (A.8a'), we find

$$c = -2(kM/r^3)\, dr \sum_{n=2}^{N} (n-1)(a/r)^n\, \Delta S(n) - (kM/r^2) \cdot \sum_{n=2}^{N} n(n-1)(a/r)^n (1/r) dr\, \Delta S(n);$$

upon considering (A.19b) this results in

$$c = -C\, \Delta P_6 (dr/r) . \qquad (A.33)$$

As a passing thought, we may mention that even if (A.8a') were not respected, the above relation could still be considered valid, simply because the increase would be merely $4C(\Delta r_o/r_o^*)(dr/r_o^*)$. With Δr_o as large as 10 m the term $4\Delta r_o/r_o^*$ still amounts to only 0.6×10^{-5}, which is two orders of magnitude smaller than the largest value of ΔP_6 figuring in (A.33). The relation (A.33) will next be used to evaluate the error in Δg_r due to the error in the radial distance r.

In order to assess the contribution of c to $d\Delta g_r$, only an approximate version of the formula (A.28) is needed, namely

$$dr \simeq dr_o + r_o \sum_{n=2}^{N} (a/r)^n \, dS(n) ,$$

which can be further modified to yield

$$dr/r \simeq (r/r_o^{*2}) \, dr_o + \sum_{n=2}^{N} (a/r)^n \, dS(n) .$$

Substituting this result into (A.33), we have

$$c \simeq C[-(r/r_o^{*2}) \, \Delta P_6 \, dr_o - \Delta P_6 \sum_{n=2}^{N} (a/r)^n \, dS(n)] .$$

Neglecting the contribution of c and taking (A.32) as the final outcome amounts to replacing $(-2 - \Delta P_6)$ and $[(n-1) - \Delta P_6]$ in the coefficients simply by -2 and (n-1), respectively. But the errors thus committed affect at most the fourth or fifth significant digit in the partial derivatives, even if the value in (A.22b) is admitted for ΔP_6.

Thinking in terms of observation equations, we realize that the "discrepancy terms" will probably be good to no more than two to three significant digits. Thus, changes in the partial derivatives due to the variable r are completely negligible and the equation (A.32) gives the desired result. This conclusion is tantamount to saying that the model for Δg_r may be regarded as linear in the parameters.

An interesting outcome in this development is the finding that if the radial distance to P were really fixed, i.e., if P were given by accurate or even perfect geocentric coordinates, the advantage of having such coordinates would not be felt in conjunction with the gravity anomaly model. This

model is incapable, for all practical purposes, of distinguishing between r considered as a known constant and r considered as a function of potential coefficients. In a plausible way, we may describe the gravity anomaly model as a linear expansion from an ellipsoid at distances of tens of meters (geoid undulations), as opposed to a nonlinear expansion from the coordinate origin at distances r.

We shall now address the problem of forming an observation equation based on the gravity anomaly model in which the equality sign will be employed instead of "≃". We notice that from the adjustment point of view, the equation (A.31') is valid for adjusted quantities (i.e., for an adjusted "observation" on the left-hand side and for adjusted parameters on the right-hand side). It can be therefore rewritten as

$$v_\Delta + (\Delta g_r)^b = (\Delta g_r)^o + d\Delta g_r \ , \qquad (A.34)$$

where

v_Δ = gravity anomaly residual,

$(\Delta g_r)^b$ = "observed" value of the observable,

$(\Delta g_r)^o$ = computed value of the observable based on the starting values (p^0) of the parameters,

$d\Delta g_r$ = correction to the above computed value, based on the corrections (dp) to p^0.

From (A.31') we have

$$(\Delta g_r)^o = C\,[\Delta P_5 - 2(r/r_o^{*2})\Delta r_o] \ . \qquad (A.35)$$

Upon using also (A.32), the equation (A.34) may be formulated as

$$v_\Delta = C\left[-2(r/r_o^{*2})\,dr_o + \sum_{n=2}^{N}(n-1)(a/r)^n\,dS(n)\right] + C\left[\Delta P_5 - 2(r/r_o^{*2})\Delta r_o\right] - (\Delta g_r)^b \ . \qquad (A.36)$$

In terms of the matrix notations, this observation equation is written as follows:

$$v_\Delta = a\,x + \ell \ , \qquad (A.37)$$

where

$$a = C\left[-2(r/r_o^{*2});\ \ldots(n-1)(a/r)^n\,P_n(\sin\bar{\phi})\ldots,\ \ldots(n-1)(a/r)^n\cos m\lambda\,P_{nm}(\sin\bar{\phi})\ldots,\ \ldots(n-1)(a/r)^n\sin m\lambda\,P_{nm}(\sin\bar{\phi})\ldots\right], \qquad (A.38a)$$

partial derivatives of the gravity anomaly model with respect to the parameters (p),

$$x = \left[dr_o;\ \ldots dC_{no}\ldots,\ \ldots dC_{nm}\ldots,\ \ldots dS_{nm}\ldots\right]^T \ , \qquad (A.38b)$$

corrections (dp) to the starting values p^0,

$$\ell = (\Delta g_r)^o - (\Delta g_r)^b \equiv C\left[\Delta P_5 - 2(r/r_o^{*2})\Delta r_o\right] - (\Delta g_r)^b \qquad (A.38c)$$

or, usually,

$$\ell = C\Delta P_5 - (\Delta g_r)^b \ , \qquad (A.38c')$$

the "discrepancy term" (the computed minus the "observed" value of the observable).

The values of r, etc., in these expressions are computed using p^0. Upon considering (A.33), even the ellipsoidal radial distance is seen to be a completely satisfactory substitution for r since it could introduce an error not exceeding 0.013 mgal even if the error in r (i.e., the geoid undulation) reached 150m.

It was shown in [B], Section 5.2, that the right-hand side of (A.31) -- although without the parameter Δr_o -- suits the quantity

$$\Delta g_r \equiv g_r - \gamma_r \tag{A.39a}$$

better than it suits the quantity

$$\Delta g \equiv g - \gamma \ . \tag{A.39b}$$

In particular, this expression gives Δg_r with only one approximation involved (the spherical approximation $\gamma \simeq kM/r^2$). The use of this expression for Δg includes not only the spherical approximation but also an additional, less serious "direction approximation", namely

$$\Delta g \simeq \Delta g_r \ . \tag{A.40}$$

The spherical approximation was shown to introduce the largest errors in certain equatorial regions, but even there the errors in Δg_r amounted to no more than 0.2 mgal.

A similar conclusion can be drawn when Δr_o is present. In fact, when deriving (A.29), we witnessed that no change in the procedure of [B] took place other than replacing T by $(T - \Delta W_o)$. The parameter ΔW_o represents the difference between the actual potential of the geoid and its "best estimate". On the average, this difference is likely to be much smaller than T,

perhaps by one or more orders of magnitude. We can thus assume that since the factor (-2/r) obtained through the spherical approximation did not introduce a significant error in Δg_r when multiplied by T, it will not do so when multiplied by $(T - \Delta W_o)$, either.

Similarly, since replacing T by $(T - \Delta W_o)$ could alter the quantities Δg_r themselves only to a very limited degree, the difference between the radial and the normal components of the gravity anomaly will remain close to that reported in [B]. The validity of the direction approximation is thus essentially unaffected by the presence of Δr_o, a conclusion similar to the one just reached with regard to the spherical approximation. This amounts to saying that in general, the usual gravity anomaly (Δg) is very nearly equal to Δg_r which, in turn, is represented quite satisfactorily by the right-hand side of (A.31).

We now describe the modeling errors in the observation equation (A.34), written also as (A.36) or (A.37). The basic equation (A.31) or (A.31') is written in terms of the spherical approximation; it thus contains the "spherical error" which is, of course, also present in any of the three forms of the observation equation just mentioned. However, it is not possible to have an "observed" Δg_r value appearing as $(\Delta g_r)^b$ in the observation equation. For this reason, the symbol Δg_r in all these expressions is replaced by Δg, which means that the "direction error" is introduced. The quantity Δg is then considered as observed, in the sense that it is derived from actual observations. On the whole, the observation equation will thus contain the spherical and the direction errors as modeling errors. Dropping the superscript "b", the observed gravity anomalies are

$$\Delta g = g - \gamma \; ; \tag{A.41}$$

the variance of Δg is equal to the variance of g. The reciprocal value of the variance becomes the weight associated with the pertinent observation equation. The geoidal point P can represent a certain area on the geoid rather than a single point, in which case the corresponding mean value of Δg and its variance have to be computed; otherwise, the two concepts are similar.

A.4 Gravity Mathematical Model

We shall now consider the force of gravity at a geoidal point P whose position is specified as in (A.1a) with r being given in (A.1b) or, more specifically, in (A.14) or (A.14'). We shall examine the change in g_r (the radial component of the gravity vector) due to the change in the parameters from their starting values. In analogy to the previous development, r will again be considered fixed and, later, a correction (c) will be added, which stems from the fact that r is in reality a function of the parameters. We have

$$dg_r(r \text{ fixed}) = C \sum_{n=2}^{N} (n+1)(a/r)^n \, dS(n) \; . \tag{A.42}$$

If we were dealing with a fictitious case where r would be indeed fixed, we would arrive at the following observation equation:

$$v\,(r\,\text{fixed}) = a\,(r\,\text{fixed})\,x + \ell \; ,$$

where

$$a\,(r\,\text{fixed}) = C\,[0;\; \ldots(n+1)(a/r)^n\, P_n(\sin\bar{\phi})\ldots,$$
$$\ldots(n+1)(a/r)^n \cos m\lambda\, P_{nm}(\sin\bar{\phi})\ldots,$$
$$\ldots(n+1)(a/r)^n \sin m\lambda\, P_{nm}(\sin\bar{\phi})\ldots],$$

and where x and ℓ would be the same as in the more realistic case now being developed.

Taking the equations (A.18d) and (A.20) into account, from (A.10') we obtain

$$c = -C(1/r)(2 + P_4 + D)\, dr \,. \tag{A.43}$$

Furthermore, the symbol F is introduced as follows:

$$F = (2 + P_4 + D)/[1 + (r_o/r)(P_2 - 3D/2)] \;; \tag{A.44}$$

one can now combine the results of (A.42) and (A.43), with dr given by (A.28), into the final formula

$$dg_r = C \{-F(1/r_o)\, dr_o + \sum_{n=2}^{N} [(n+1) - F(r_o/r)](a/r)^n\, dS(n)\} \,. \tag{A.45}$$

We are now in a position to form an observation equation at P. Noting that the equation (A.10') is valid for adjusted quantities, we rewrite it as

$$v + (g_r)^b = (g_r)^o + dg_r \,, \tag{A.46}$$

where the symbols could be described in a complete analogy to the description which followed (A.34) in dealing with the gravity anomaly model. Upon realizing from (A.10') that

$$(g_r)^o = C(1 + P_3 - D) \,, \tag{A.46'}$$

where the right-hand side is evaluated through p^0, the equation (A.46) may be formulated as

$$v = C \{-F(1/r_o)\, dr_o + \sum_{n=2}^{N} [(n+1) - F(r_o/r)](a/r)^n\, dS(n)\} + C(1 + P_3 - D) - (g_r)^b \,. \tag{A.47}$$

In terms of the matrix notations, this observation equation is written as follows:

$$v = ax + \ell \ , \tag{A.48}$$

where

$$a = C \{-F(1/r_o); \ldots[(n+1) - F(r_o/r)](a/r)^n P_n(\sin\bar{\phi})\ldots,$$
$$\ldots[(n+1) - F(r_o/r)](a/r)^n \cos m\lambda \, P_{nm}(\sin\bar{\phi})\ldots,$$
$$\ldots[(n+1) - F(r_o/r)](a/r)^n \sin m\lambda \, P_{nm}(\sin\bar{\phi})\ldots\}, \tag{A.49a}$$

$$x = [dr_o; \ldots dC_{no}\ldots, \ldots dC_{nm}\ldots, \ldots dS_{nm}\ldots]^T \ , \tag{A.49b}$$

$$\ell = (g_r)^o - (g_r)^b \equiv C(1 + P_3 - D) - (g_r)^b \ . \tag{A.49c}$$

The description would again be similar to that presented in connection with the gravity anomaly observation equation.

However, the replacement of r by the ellipsoidal radial distance is no longer possible. From the equation (A.43) giving the error in g_r as a function of an error in r (the latter is represented here by geoid undulations), we conclude that the geoid undulation of 150m would result in an error of approximately 45 mgal. To produce errors smaller than 0.03 mgal, r should be computed with better than a 10 cm accuracy. One way to achieve this would be to proceed to an iterative solution of (A.14'). However, an algorithm developed in Chapter 2 of the present report allows the computation of r to a sub-centimeter accuracy without iterations.

We now address the problem of obtaining the "observation" $(g_r)^b$ in practice. In this task use is made of the direction approximation (A.40) written in the form

$$g_r - g \simeq \gamma_r - \gamma \ .$$

In Appendix 2 of [B], the following formula was explicitly derived:

$$\gamma - \gamma_r = f(a, J_2, \ldots, \bar{\phi}) \ ,$$

where $J_2 = -C^*_{20}$. The direction approximation is then implied in the statement

$$g_r \simeq g - f(a, J_2, \ldots, \bar{\phi}) \ . \qquad \text{(A.50)}$$

The "observation" $(g_r)^b$ is thus represented by the right-hand side of (A.50), containing only the direction error. Unlike in the gravity anomaly model, no spherical error is present. The variance entering the adjustment is that of g. It is now apparent that the direction error stems from treating only one component of g (namely g_r). The other two components $(g_{\bar{\phi}}, g_\lambda)$ are assumed to be related to this, by far the largest, component in exactly the same manner as they are in the normal gravity field. However, the direction error is insignificant and does not warrant complicating the mathematical model; although it is in general much smaller than the spherical error, its removal would complicate the gravity model substantially more than the removal of the spherical error has done.

The observation equation described in (A.48), (A.49) proves also useful when one considers deducing gravity observations from gravity anomalies, namely

$$g = \Delta g + \gamma \ . \qquad \text{(A.51a)}$$

This together with (A.50) yields

$$g_r \simeq \Delta g + \gamma - f(a, J_2, \ldots, \bar{\phi}) \; ; \qquad \text{(A.51b)}$$

the variance entering the adjustment is that of Δg. Other than replacing $(g_r)^b$ by the right-hand side of (A.51b), the adjustment algorithm is exactly the same as if the gravity model were employed from the beginning.

A.5 Comparison of the Gravity and the Gravity Anomaly Models

In theory, the comparison in this section could be made at the level of the formulas for Δg_r and g_r, provided γ_r is subtracted from both sides of the equation giving the latter. The difficulty of such a task becomes apparent already at the stage of listing all the equations involved. On one hand, we have Δg_r appearing in (A.31'). On the other hand, we would have to consider g_r given in (A.10') and r given in (A.14'); furthermore, γ_r involves the equation (A.12) with $r = \tilde{r}$, and $\tilde{r}$ involves (A.16). The spherical approximation was used in deriving (A.31'), but not in the other listed equations.

Fortunately, a different approach can be undertaken in order to prove a basic equivalence between the gravity and the gravity anomaly models. Perhaps the easiest solution in such a task is to perform the comparison at the level of observation equations. In fact, the basic idea behind our approach consists in showing that the gravity observation equation (A.47) is identical, for most practical purposes, to the gravity anomaly observation equation (A.36). A perfect identity cannot be produced because of the spherical approximation implicitly present in the gravity anomaly model. We begin our demonstration by adding and subtracting γ_r when treating the right-hand side of (A.47). From the definition (A.39a), i.e., from

$$\Delta g_r = g_r - \gamma_r \qquad \text{(A.52a)}$$

we have

$$(\Delta g_r)^b = (g_r)^b - \gamma_r \ , \tag{A.52b}$$

so that (A.47) now reads

$$v = C \{-F(1/r_o)\, dr_o + \sum_{n=2}^{N} [(n+1) - F(r_o/r)]\, (a/r)^n\, dS(n)\}$$
$$+ [C(1 + P_3 - D) - \gamma_r] - (\Delta g_r)^b \ . \tag{A.53}$$

To show that (A.53) is essentially (A.36) requires a proof that the respective coefficients of dr_o and of $dS(n)$ are very nearly equal and, further, that

$$C(1 + P_3 - D) - \gamma_r = C\, [\Delta P_5 - 2(r/r_o^{*2})\Delta r_o] \tag{A.54a}$$

holds within negligible margins; upon considering (A.35) and (A.46'), the above equation can be also written as

$$(g_r)^o - \gamma_r = (\Delta g_r)^o \ . \tag{A.54b}$$

Due to (A.8a'), Δr_o in our demonstrations will be zero (the superscript "o" has been omitted). The proof at the level of observation equations is relatively easy because it involves merely numerical equivalences between values computed via the parameters p^o. This enables us to make various simplifications and to immediately evaluate their numerical effect.

Having a practical purpose in mind, we can afford to demonstrate (A.54) within a few tenths of a milligal since usually $(\Delta g_r)^b$ will not be known to a much better accuracy. Thus (A.54) will be good to one or two significant digits most of the time. On the other hand, the equivalence

between the respective coefficients of the observation equations will be proved to at least two significant digits all of the time. In doing so, we shall be considering the worst possible errors which could occur upon equating the coefficients or the constant terms between the two types of observation equations. Both types of errors (i.e., in the coefficients and in the constant terms) will be seen to be the most important in certain equatorial regions. Accordingly, we shall pay most attention to the errors associated with $\overline{\phi}=0$, even though other types of knowledge (e.g., that of geoid undulations) would be needed for a rigorous proof that maximum errors in all cases indeed occur at or near the equator.

The above proof of equivalence is mostly of practical interest in that it indicates that the gravity model must lead to essentially the same results as obtained through the gravity anomaly model. This applies of course not only for the parameters, but also for the adjusted observations, predictions, and the variance-covariance propagation. The possible small differences stem mainly from the spherical approximation used in the gravity anomaly model. As we have seen earlier, the error caused by neglecting certain reference coefficients is almost always negligible. Additional approximations sometimes made in practice can be eliminated and need not be considered here. Our demonstration consists of two steps. We first show the equivalence between the corresponding coefficients in the observation equations, and then the equivalence of the discrepancy terms by demonstrating (A.54b).

A.5.1 Equivalence of the Coefficients

We observe that (r_o/r) in the formula (A.44) for F may safely be replaced by unity, the errors thus introduced being two orders of magnitude smaller than those we are concerned with. We first make the following approximation:

$$2 + P_4 + D \simeq 2;$$

the error which will later participate in the worst total effect occurs at a certain point whose latitude is $\bar{\phi} = 0$ (we use the original minus the simplified value in defining such errors). This error amounts to +0.0104 according to the specific numerical results listed in (A.21c) and (A.23b), and it represents 0.5% of the above simplified value (i.e., 2). Further, we make the following approximation in the denominator of F:

$$1 + P_2 - 3D/2 \simeq 1,$$

which could simlarly introduce a -0.4% error at $\bar{\phi} = 0$.

In accordance with the above, the approximation

$$F \simeq 2 \qquad \text{(A.55a)}$$

could introduce, in the worst case, a +0.9% error in this value (numerically, it would be +0.018). The simplification

$$r_o/r \equiv 1/(1 + P_1 + \tfrac{1}{2}D) \simeq 1$$

results in a maximum error of -0.2% for $\bar{\phi} = 0$. Consequently, the expression

$$F(1/r_o) \simeq 2r/r_o^{*2} , \qquad \text{(A.55b)}$$

or (since $r_o = r_o^*$ is assumed throughout when dealing with numerical values),

$$F(r_o/r) \simeq 2 \qquad \text{(A.55c)}$$

could introduce, in the worst possible case, a +0.7% error of the term's magnitude; as a matter of interest, at the pole we would have -0.35% instead. We can write (A.55c) as

$$n + 1 - F(r_o/r) \simeq n - 1 \ , \qquad \text{(A.55d)}$$

which implies a numerical error of -0.014. If $n = 2$, this represents a -1.4% error of the term's magnitude, which symbolizes the worst possible conditions not only with respect to the geographic position, but also with respect to n. If $n = 3, n = 4$, etc., the worst possible errors in the pertinent coefficients would be -0.7%, -0.5%, etc., respectively.

In order to verify the above error estimates, the values of $F(r_o/r)$ were computed in a $5^0 \times 5^0$ geographic grid, upon utilizing the (20,20) set of potential coefficients mentioned earlier. The largest deviations from the approximate values in (A.55a), (A.55c), and (A.55d) occurred at the location $\overline{\phi} = 0$, $\lambda = 5^0$. The exact values at that location were

$$\begin{aligned} F &= 2.0188, \\ F(r_o/r) &= 2.0142, \\ n + 1 - F(r_o/r) &= n - 1 - 0.0142. \end{aligned}$$

Accordingly, the relative errors are respectively +0.94%, +0.71%, and -1.42% in case $n = 2$ (if $n = 3$ this last error is -0.71%, if $n = 4$ it is -0.47%, etc.). The numerical results thus agree very well with the approximate semi-analytical estimates as presented in the last paragraph.

Substituting the results (A.55b) and (A.55d) in (A.53), we have to a good approximation:

$$v = C\left[-2(r/r_o^{*2})\, dr_o + \sum_{n=2}^{N} (n-1)(a/r)^n\, dS(n)\right] + \left[C(1 + P_3 - D) - \gamma_r\right] - (\Delta g_r)^b \,. \qquad (A.56)$$

The first part of this expression has already the desired form (A.36), which amounts to writing

$$dg_r = d\Delta g_r \,.$$

Although it represents merely an intermediate step in the current equivalence demonstration, the equation (A.56) is an interesting result in its own right. First of all, it may be mentioned that $-F(1/r_o)$ in (A.53) could have been replaced by $-2/r_o$, in which case the first term on the right-hand side of (A.56) would be simplified to read $-(2/r_o)\, dr_o$. According to (A.55a), the error in this new coefficient $-(2/r_o)$ could reach +0.9% of its magnitude rather than +0.7% associated with (A.55b). The worst possible error in any other coefficient would be -1.4% and it could occur only for $n = 2$. The equation (A.56) is in fact a mixture of the gravity anomaly model (the part comprising the coefficients) and the gravity model (the constant term). If the constant term is written in its original form as in (A.47) and the just discussed simplification is taken into account, (A.56) may be rewritten as

$$v = C\left[-(2/r_o)\, dr_o + \sum_{n=2}^{N} (n-1)(a/r)^n\, dS(n)\right] + \left[C(1 + P_3 - D) - (g_r)^b\right]. \qquad (A.56')$$

This observation equation embodies quite well both the accuracy of the gravity model and the simplicity of the gravity anomaly model.

In the matrix notations, (A.56') may be presented as follows:

$$v = ax + \ell \ ,$$

$$a = C\,[-(2/r_o);\ \ldots(n-1)(a/r)^n\ P_n(\sin\bar{\phi})\ldots,$$

$$\ldots(n-1)(a/r)^n \cos m\lambda\ P_{nm}(\sin\bar{\phi})\ldots,$$

$$\ldots(n-1)(a/r)^n \sin m\lambda\ P_{nm}(\sin\bar{\phi})\ \ldots]\ ,$$

$$x = [dr_o;\ \ldots dC_{no}\ldots,\ \ldots dC_{nm}\ldots,\ \ldots dS_{nm}\ldots]^T\ ,$$

$$\ell = (g_r)^o - (g_r)^b \equiv C(1 + P_3 - D) - (g_r)^b.$$

The formula for "a" corresponds to (A.38a) and "ℓ" is (A.49c). The "observation" $(g_r)^b$ can be obtained either as in (A.50) or as in (A.51b).

The parameter r_o and often also C_{20} (and perhaps several other coefficients) are in practice held fixed or weighted, especially if the data are insufficient for their reliable determination. The parameters C_{21} and S_{21} can always be held fixed at the zero value in accordance with (A.2b). If C_{22} and S_{22} are also held fixed or weighted, the coefficients in (A.56') whose errors could have a noticeable influence on the adjustment start with $n = 3$ (or a higher degree). In this situation, the worst possible error would be -0.7% of the coefficients' magnitude for $n = 3$, as opposed to -1.4% for $n = 2$. We can say that in general all the coefficients in the above observation equation will be reliable to at least two significant digits, the accuracy improving drastically with the increasing n. On the other hand, in the presence of a

"good" set of initial values of the parameters, in the sense that the corrections to these values as obtained from an adjustment are relatively small, the constant terms "ℓ" will be in general small. In fact, they may reach, for the most part, only a few milligals and thus only one or two significant digits may be meaningful. But the reliability of the coefficients in the observation equations represented by (A.56') has just been shown to be in general better; they are thus completely satisfactory when used in conjunction with the "gravity-type" constant terms. These new coefficients can also accelerate the computation of predicted values and the corresponding variances-covariances.

In conclusion of this discussion, we may state that the observation equation (A.56') is essentially the gravity anomaly observation equation except for the constant term. This term corresponds to the gravity model and as such, it embodies the improvement in accuracy resulting from the removal of the spherical approximation. Returning now to (A.56), we may conclude that the first part of our demonstration has been completed.

A.5.2 Equivalence of the Constant Terms

We first restate the equation (A.10) for a geoidal point (P) in which the initial values p^0 are utilized, obtaining

$$(g_r)^0 = C\left[1 + \sum_{n=2}^{N} (n+1)(a/r)^n \sum_{m=0}^{n} (C^0_{nm} \cos m\lambda + S^0_{nm} \sin m\bar{\phi}) \cdot P_{nm}(\sin\bar{\phi})\right] - \omega^2 r \cos^2\bar{\phi} \,, \qquad (A.57)$$

where r is computed from (A.14) as

$$r = r_o^* \left[1 + \sum_{n=2}^{N} (a/r)^n \sum_{m=0}^{n} (C^o_{nm} \cos m\lambda + S^o_{nm} \sin m\lambda) \cdot P_{nm}(\sin\bar{\phi}) + \tfrac{1}{2}D\right] . \qquad (A.58)$$

Although in practice the equations (A.8a') and (A.8b') will both be used most of the time, we note that only (A.8a') results in important simplifications in the derivations. Accordingly we adopt (the superscript "o" is omitted)

$$r_o = r_o^* , \qquad \Delta r_o = 0 , \qquad (A.59)$$

as can be already witnessed from (A.58).

For the sake of simplicity, we now prefer to restate (A.12) for the corresponding ellipsoidal point ($\tilde{P}$) as follows:

$$\gamma_r = \tilde{C}\left[1 + 3(a/\tilde{r})^2 C^*_{20} P_2(\sin\bar{\phi}) + 5(a/\tilde{r})^4 C^*_{40} P_4(\sin\bar{\phi})\right] - \omega^2\tilde{r}\cos^2\bar{\phi} , \qquad (A.60)$$

where, from (A.16),

$$\tilde{r} = r_o^*\left[1 + (a/\tilde{r})^2 C^*_{20} P_2(\sin\bar{\phi}) + (a/\tilde{r})^4 C^*_{40} P_4(\sin\bar{\phi}) + \tfrac{1}{2}\tilde{D}\right] ; \qquad (A.61)$$

in these formulas, the notations $\tilde{C}$ and $\tilde{D}$ have been introduced in analogy to (A.20). Finally, we present the formula for $(\Delta g_r)^0$ needed in proving (A.54b). Taking advantage of (A.59), we write it from (A.31) as

$$(\Delta g_r)^0 \simeq C \sum_{n=2}^{N} (n-1)(a/r)^n \sum_{m=0}^{n} (\Delta C^o_{nm} \cos m\lambda + \Delta S^o_{nm} \sin m\lambda) \cdot P_{nm}(\sin\bar{\phi}) . \qquad (A.62)$$

We next define the quantity Δr, similar in nature to the geoid undulation (here we refer to the geoid as approximated through p^0):

$$\Delta r = r - \tilde{r} \ . \tag{A.63}$$

The change in $\tilde{C}$ due to Δr is

$$dC \equiv C - \tilde{C} = -2\tilde{C}(\Delta r/\tilde{r}) \ . \tag{A.64a}$$

Using $\tilde{C} = 980$ gal and some mean value for $\tilde{r}$, this may be written as

$$dC = -0.31 \ \Delta r \ , \tag{A.64b}$$

where dC is in milligals if Δr is meters, and it corresponds to an average free-air reduction. Similarly,

$$dD \equiv D - \tilde{D} = 3\tilde{D}(\Delta r/\tilde{r}).$$

Adopting $\tilde{D} \simeq 0.0034 \cos^2\overline{\phi}$, we have

$$dD = +0.0016 \times 10^{-6} \cos^2\overline{\phi} \ \Delta r \ , \tag{A.65}$$

where dD is unitless and Δr is in meters.

It is apparent from (A.57) and (A.60) that when we attempt to form $(g_r)^0 - \gamma_r$, the first obstacle we encounter is the presence of two different types of radial distances, r and $\tilde{r}$. Our immediate task, then, is to find a formula for Δr in terms of r and to evaluate the error introduced. When forming the difference between (A.58) and (A.61), one obtains

$$\Delta r \simeq r_o^* \Big[\sum_{n=2}^{N} (a/r)^n \sum_{m=0}^{n} (C_{nm}^0 \cos m\lambda + S_{nm}^0 \sin m\lambda) P_{nm}(\sin\overline{\phi}) \\ -(a/\tilde{r})^2 \, C_{20}^* \, P_2(\sin\overline{\phi}) - (a/\tilde{r})^4 \, C_{40}^* \, P_4(\sin\overline{\phi}) \Big] . \tag{A.66}$$

The error associated with this approximation is

$$\varepsilon_1 = \tfrac{1}{2}\, r_o^*\, dD$$

or, upon taking (A.65) into account,

$$\varepsilon_1 = +0.0051 \cos^2\bar{\phi}\ \Delta r, \qquad \varepsilon_1 = +0.0051\ \Delta r \ldots \bar{\phi} = 0,$$

both ε_1 and Δr being in meters.

In the last two terms of (A.66), $\tilde{r}$ has to be replaced by r which introduces another error. To a sufficient accuracy we have

$$(a/\tilde{r})^n = (a/r)^n + n(\Delta r/\tilde{r})\ . \tag{A.67}$$

If we now introduce the simplification

$$-r_o^*(a/\tilde{r})^n\ C_{no}^*\ P_n(\sin\bar{\phi}) \simeq -r_o^*(a/r)^n\ C_{no}^*\ P_n(\sin\bar{\phi}),$$

we have introduced the error

$$\varepsilon_2 = -n\ \Delta r\ C_{no}^*\ P_n(\sin\bar{\phi}).$$

Adopting (from GRS 1967) the reference value $C_{20}^* = -0.0010827$, for $n = 2$ we have

$$\varepsilon_2 = +0.0022\ P_2(\sin\bar{\phi})\ \Delta r, \qquad \varepsilon_2 = -0.0011\ \Delta r\ \ldots\bar{\phi} = 0\ .$$

The error associated with $n = 4$ is completely negligible (one would have $C_{40}^* = +0.000\ 0024$ and, for $\bar{\phi} = 0$ and Δr as large as 150m, $|\varepsilon_2| < 0.6$ mm).

Multiplying (A.66) by r/r_o^*, on the left-hand side we have $\Delta r + \Delta r(r - r_o^*)/r_o^*$ while on the right-hand side r_o^* has been replaced by r (previously $\tilde{r}$ had been replaced by r). The equation (A.66) may thus be written as

$$\Delta r \simeq r \sum_{n=2}^{N} (a/r)^n \sum_{m=0}^{n} (\Delta C^0_{nm} \cos m\lambda + \Delta S^0_{nm} \sin m\lambda) P_{nm}(\sin\bar{\phi}), \tag{A.68}$$

while the additional error just introduced is

$$\varepsilon_3 = -\Delta r(r - r^*_o)/r^*_o, \qquad \varepsilon_3 = -0.0023\ \Delta r \ \ldots\ \bar{\phi} = 0\ ,$$

where (from GRS 1967) $a = 6{,}378{,}160\,m$ and $r^*_o = 6{,}363{,}696\,m$ have been utilized. The total error is

$$\varepsilon = \varepsilon_1 + \varepsilon_2 + \varepsilon_3\ ,$$

and thus

$$\varepsilon = +0.0017\ \Delta r \ \ldots\ \bar{\phi} = 0\ . \tag{A.69}$$

This is not the worst possible error in Δr per se (at the poles one would find similarly $+0.0033\ \Delta r$), but it will contribute toward the worst error (ε') in the quantity $(g_r)^0 - \gamma_r$. If we accept $\Delta r = 150\,m$, which is almost 50% higher than the magnitude of the largest geoid undulations usually encountered, the error ε in (A.69) will amount to 0.26 m.

We now subtract (A.60) from (A.57), yielding

$$(g_r)^0 - \gamma_r \simeq A_1 + A_2 + A_3\ , \tag{A.70}$$

where

$$A_1 = C - \tilde{C} \equiv dC\ , \tag{A.71a}$$

$$A_2 = C \sum_{n=2}^{N} (n+1)(a/r)^n \sum_{m=0}^{n} (C^0_{nm} \cos m\lambda + S^0_{nm} \sin m\lambda) \times P_{nm}(\sin\bar{\phi})\ , \tag{A.71b}$$

$$A_3 = -\tilde{C}\,[3(a/\tilde{r})^2\, C^*_{20}\, P_2(\sin\bar{\phi}) + 5(a/\tilde{r})^4\, C^*_{40}\, P_4(\sin\bar{\phi})]\ , \tag{A.71c}$$

and where the following error has been introduced:

$$\varepsilon_1' = -\omega^2 \cos^2\bar{\phi}\ \Delta r\ ;$$

considering the numerical value of ω (about 0.000 073 rad/sec), we observe that

$$\varepsilon_1' = -0.00053 \cos^2\bar{\phi}\ \Delta r, \qquad \varepsilon_1' = -0.00053\ \Delta r \ \ldots\ \bar{\phi} = 0.$$

It is clear that in a linearized form, dC given by (A.64a) could have been equally well written with C,r replacing $\tilde{C}$,$\tilde{r}$. Thus we have

$$A_1 = -2C(\Delta r/r)$$

or, making use of (A.68),

$$A_1 = -2C \sum_{n=2}^{N} (a/r)^n \sum_{m=0}^{n} (\Delta C^0_{nm} \cos m\lambda + \Delta S^0_{nm} \sin m\lambda)\, P_{nm}(\sin\bar{\phi})\ . \tag{A.71a'}$$

However, this expression contains an error (ε_2') introduced through the error (ε) in Δr depicted in (A.69). From the approximate form of A_1 exhibited in (A.64b), we conclude that the contribution of ε in this case is not negligible, its order of magnitude being comparable with the other error terms containing directly Δr. We deduce that

$$\varepsilon_2' = -0.31\,\varepsilon, \qquad \varepsilon_2' = -0.000\ 53\ \Delta r \ \ldots\ \bar{\phi} = 0,$$

which is of the same form as ε_1'.

The expression for A_2 in (A.71b) needs no modification. On the other hand, the quantities $\tilde{r}$ in (A.71c) will have to be "converted" into r. First, using (A.67), we write

$$-\tilde{C}(n+1)(a/\tilde{r})^n\ C^*_{no}\ P_n(\sin\bar{\phi}) \simeq -\tilde{C}(n+1)(a/r)^n\ C^*_{no}\ P_n(\sin\bar{\phi})\ ,$$

where the error is

$$\varepsilon_3' = -\tilde{C}(n+1)n\ (\Delta\tilde{r}/r)C^*_{no}\ P_n(\sin\bar{\phi})\ .$$

For $n = 2$, this yields

$$\varepsilon_3' = +0.0010\ P_2(\sin\bar{\phi})\ \Delta r, \qquad \varepsilon_3' = -0.00050\ \Delta r\ \ldots\bar{\phi} = 0\ .$$

(Associated with C^*_{40}, we would have completely negligible errors.)

We still have to evaluate the additional error caused by replacing also $\tilde{C}$ in (A.71c) by C. Denoting the quantity in the brackets as B, one obtains

$$B \simeq -0.0032\ P_2(\sin\bar{\phi}) + 0.000\ 012\ P_4(\sin\bar{\phi})$$

and

$$B \simeq +0.0016\ \ldots\ \bar{\phi} = 0\ .$$

The above mentioned error is then given as

$$\varepsilon_4' = +dC\,B\,, \qquad \varepsilon_4' = -0.000\ 50\ \Delta r\ \ldots\ \bar{\phi} = 0\,,$$

where (A.64b) has been taken into account. We have thus determined that A_3 may be written in the form

$$A_3 \simeq -C\ [3(a/r)^2\ C^*_{20}\ P_2(\sin\bar{\phi}) + 5(a/r)^4\ C^*_{40}\ P_4(\sin\bar{\phi})]\,, \qquad \text{(A.71c')}$$

and we have shown what it entails in terms of errors (see $\varepsilon_3' + \varepsilon_4'$).

If we now consider the equations (A.71b) and (A.71c'), we can immediately write

$$A_2 + A_3 = C \sum_{n=2}^{N} (n+1)(a/r)^n \sum_{m=0}^{n} (\Delta C^0_{nm} \cos m\lambda + \Delta S^0_{nm} \sin m\lambda) \cdot P_{nm}(\sin\bar{\phi})\ . \qquad \text{(A.72)}$$

By adding A_1 from (A.71a') to this result in agreement with (A.70), the new right-hand side is the same as the right-hand side of (A.62). It then follows that

$$(g_r)^0 - \gamma_r \simeq (\Delta g_r)^0 \; ; \tag{A.73}$$

this is in fact (A.54b) which we set out to prove. This expression is subject to the error ε', namely

$$\varepsilon' = \varepsilon'_1 + \varepsilon'_2 + \varepsilon'_3 + \varepsilon'_4 \; ,$$

and thus to

$$\varepsilon' = -0.0021 \; \Delta r \; \ldots \; \bar{\phi} = 0 \; . \tag{A.74}$$

As a matter of interest, we mention that the counterpart of (A.74) for the poles would yield +0.0010 Δr, due to cancellations. If we accepted $\Delta r = 150$ m, the error ε' in (A.74) would reach -0.31 mgal.

To make a realistic evaluation of the error ε' at a certain point on the equator, the approximate geoid undulation (N) as defined in (A.63) would have to be computed at that location. However, this can be done with the aid of the values already given. The approximate formula for N reads

$$N \simeq R \sum_{n=2}^{N} (a/r)^n \, \Delta S(n) \; ,$$

where the notation (A.17b) has been utilized and where R represents the earth's mean radius. Upon approximating $(a/r)^n$ by unity, we can further write

$$N \simeq R \left[\sum_{n=2}^{N} S(n) - C^*_{20} \, P_2(\sin\bar{\phi}) - C^*_{40} \, P_4(\sin\bar{\phi}) \right] .$$

The summation in this expression corresponds to P_1 as presented in (A.21a) and denoted now as P_1'.

On the equator, we have

$$N \simeq R(P_1' - 0.000\ 5422) \ . \qquad \text{(A.75a)}$$

The value presented in (A.21a) for $\bar{\phi} = 0$ represents a maximum which occurred at $\lambda = 150^0$; its more exact numerical value is

$$P_1' = +0.000\ 5520 \ \ldots \ \lambda = 150^0 \ . \qquad \text{(A.75b)}$$

The minimum value for $\bar{\phi} = 0$ has not yet been presented. However, this value will lead to the largest magnitude of N in this investigation; thus we list

$$P_1' = +0.000\ 5278 \ \ldots \ \lambda = 75^0 \ . \qquad \text{(A.75c)}$$

Upon considering (A.75b) and (A.75c), respectively, the equation (A.75a) yields

$$N \simeq +62\text{m} \ \ldots \ \bar{\phi} = 0, \ \lambda = 150^0 \ , \qquad \text{(A.76a)}$$

$$N \simeq -92\text{m} \ \ldots \ \bar{\phi} = 0, \ \lambda = 75^0 \ . \qquad \text{(A.76b)}$$

The symbols N in (A.76) represent Δr to be used in (A.74). Carrying out the indicated multiplications, we find

$$\varepsilon' = -0.13 \text{ mgal} \ \ldots \ \bar{\phi} = 0, \ \lambda = 150^0 \ , \qquad \text{(A.77a)}$$

$$\varepsilon' = +0.19 \text{ mgal} \ \ldots \ \bar{\phi} = 0, \ \lambda = 75^0 \ . \qquad \text{(A.77b)}$$

The notation ε' was shown to represent the error associated with (A.73), namely

$$\varepsilon' \equiv [(g_r)^0 - \gamma_r] - (\Delta g_r)^0 \ . \qquad \text{(A.78)}$$

According to (A.39a) and the discussion which followed, the expression

$$g_r - \gamma_r \simeq \Delta g_r$$

involves the spherical approximation and $(g_r - \gamma_r) - \Delta g_r$ is the spherical error. The quantity ε' in (A.78) depicts an estimate of this error obtained in four steps (see ε_1', ε_2', ε_3', ε_4') and involving a given set of potential coefficients.

We have seen that the set of potential coefficients adopted in this study is identical to the set used in Section 5.4 of [B], where the spherical error was computed in a $5^0 \times 5^0$ geographic grid. The largest spherical error was found at the location $\bar{\phi} = 0$, $\lambda = 75^0$; also listed was the spherical error for $\bar{\phi} = 0$, $\lambda = 150^0$ and many other locations. It was computed rigorously as the right-hand side of (A.78) or, in terms of (A.54a), as $[C(1 + P_3 - D) - \gamma_r] - C\Delta P_5$. The following pertinent results were obtained (the spherical error was denoted as Δ_2):

$$\Delta_2 = -0.130\,\text{mgal} \quad \ldots \quad \phi = 0, \quad \lambda = 150^0 \,,$$

$$\Delta_2 = +0.194\,\text{mgal} \quad \ldots \quad \phi = 0, \quad \lambda = 75^0 \,.$$

These results agree very well with our semi-analytic procedure in which, at every step, simplifications have been introduced and only two significant digits have been retained. The rigorous results thus confirm and verify the presented procedure.

Finally, we insert a remark concerned with the number of potential coefficients which have participated in the numerical outcomes. In our evaluations, the truncation has occurred at the degree and order $n = 20$,

implying that 441 C's and S's have been present (the five "forbidden" potential coefficients are included in this number). According to the equation (7) in [Rapp,1970], n = 20 roughly corresponds to mean anomalies in $9^0 \times 9^0$ blocks. The extreme values of $(\Delta g_r)^0$ computed from (A.35) with the present set of C's and S's have been found to be in the vicinity of ± 41 mgal. On the other hand, the maximum value among the mean free-air anomalies as presented in [Heiskanen and Moritz,1967], Figure 3-18, is 49 mgal. However, the figure depicts the blocks of dimensions $5^0 \times 5^0$, which corresponds to n = 36 and thus to 1369 C's and S's. Although this is more than three times the number of potential coefficients we have been using, the orders of magnitude in gravity anomalies are compatible. This indicates that the equivalences demonstrated during the current analysis would have been obtained to about the same degree of accuracy also if a larger set of potential coefficients had been utilized.

A.6 Conclusion

The most important theoretical outcome of this analysis has been the proof of a basic equivalence between the gravity anomaly model and the gravity model, both considered at the same geoidal point and expanded in terms of spherical harmonics. This equivalence has been demonstrated at the level of observation equations, i.e., after a linearization process.

In deriving an observation equation for the gravity model, it would have been insufficient to differentiate the basic equation only with respect to the parameters (potential coefficients) appearing in it explicitly. Such a procedure would be justified only if the radial distance from the geocenter to the point considered were given independently of any representation of the geoid and hence of the parameters that define it. Accordingly, it has been further necessary to differentiate the radial distance with respect to the parameters and to include the outcome of this operation in the observation equation.

The most tedious part of the comparison between the two models (gravity anomaly versus gravity) has been the comparison of the constant terms in the pertinent observation equations. This task has been carried out in a semi-analytical fashion; the needed numerical values have been obtained with the aid of a reasonable set of potential coefficients truncated at the degree and order (20,20). The agreement between the two models has

been found to be very close. Furthermore, a by-product which surfaced during the comparison of the constant terms points to a remarkable agreement between the results in the present study and those appearing in [Blaha, 1977], in that the constant terms differ, almost exactly, by the amount computed in Section 5.4 of this reference under the name of "spherical error". The spherical error, reaching at most 0.2 mgal, is present in the gravity anomaly model but not in the gravity model. Both models contain another, less important error called "direction error". The above agreement confirms and illustrates the previous theoretical development as well as numerical results; at the same time, it serves as a verification in the present comparison.

Finally, we list the formulas representing the investigated models at the level of observation equations. The gravity anomaly model appears in (A.36) or, in the matrix notations, in (A.37), (A.38). This model contains both the spherical and the direction errors.

The gravity model is presented in (A.47) or, in the matrix notations, in (A.48), (A.49). It contains only the direction error. If the data consisted of gravity values referred to the geoid, the formula (A.50) would be used in conjunction with (A.49c) giving the constant term. If, on the other hand, the data set is represented by gravity anomalies as is customary in practice, the formula to be used in conjunction with (A.49c) is (A.51b). In this way, the gravity anomaly model can be transformed into the more accurate gravity model, although the gain in accuracy may not be significant in practice and may be offset by a somewhat larger amount of computations. However, the refinement of the gravity anomaly model has been achieved, at least in theory.

As a result of certain simplifications in the gravity model, a "mixed" model has been derived. It appears in (A.56') and in the equations below it. Here again, the constant term can be computed either with the aid of (A.50) or, more usually, with the aid of (A.51b), depending on the kind of data to be adjusted. The partial derivatives correspond to the gravity anomaly model, while the constant terms are computed as in the gravity model. The "mixed" model thus approaches the gravity model in accuracy and the gravity anomaly model in simplicity, and could be viewed as a compromise between the two.

REFERENCES

Arur, M.G., Experiments for Improved Positioning by Means of Integrated Doppler Satellite Observations and the NNSS Broadcast Ephemeris. Department of Geodetic Science, Report No. 258, The Ohio State University, Columbus, 1977.

Blaha, G., The Combination of Gravity and Satellite Altimetry Data for Determining the Geoid Surface. Report of DBA Systems, Inc.; AFCRL Report No. 75-0347, Air Force Cambridge Research Laboratories, Hanscom AFB, Massachusetts, 1975.

Blaha, G., Refinements in the Combined Adjustment of Satellite Altimetry and Gravity Anomaly Data. Report of DBA Systems, Inc.; AFGL Technical Report No. 77-0169, Air Force Geophysics Laboratory, Hanscom AFB, Massachusetts, 1977.

Blaha, G., "Refinement of the Short Arc Satellite Altimetry Adjustment Model". Paper published in Bulletin Géodésique, Vol. 51, No. 1, Bureau Central de l'Association Internationale de Géodésie, Paris, France, 1977'.

Blaha, G., "An Accurate Non-Iterative Algorithm for Computing the Length of the Position Vector to a Subsatellite Point". Paper published in Bulletin Géodésique, Vol. 52, No. 3, Bureau Central de l'Association Internationale de Géodésie, Paris, France, 1978.

Brown, D.C., Review of Current Geodetic Satellite Programs and Recommendations for Future Programs. Report for NASA Headquarters, Contract No. NASW-1469; 1967.

Brown, D.C., Investigation of the Feasibility of a Short Arc Reduction of Satellite Altimetry for Determination of the Oceanic Geoid. Report No. AFCRL-TR-73-0520, Air Force Cambridge Research Laboratories (LW), Bedford, Massachusetts, 1973.

Heiskanen, W.A. and H. Moritz, Physical Geodesy. W.H. Freeman and Co., San Francisco, 1967.

Hotine, M., Mathematical Geodesy. Monogr. Ser., Vol. 2, Environ. Sci. Serv. Admin., Washington, D.C., 1969.

Marsh, J.G., M.J. Munteanu, T.V. Martin, J.J. McCarthy, P.S. Chovitz, "Estimation of the Mean Sea Surface in the North Atlantic Using GEOS-3 Altimeter Data". Paper presented at the Spring Meeting of the American Geophysical Union, Miami Beach, Florida, April 17-21, 1978.

Moritz, H., "Fundamental Geodetic Constants", pp. 129-144 in Contributions of the Graz Group to the XVI General Assembly of IUGG/IAG in Grenoble 1975, edited by P. Meissl, H. Moritz, K. Rinner; Technical University at Graz, Institute of Physical Geodesy, Graz, Austria, 1975.

Needham, P.E., The Formation and Evaluation of Detailed Geopotential Models Based on Point Masses. Department of Geodetic Science, Report No. 149, The Ohio State University, Columbus, 1970.

Rapp, R.H., The Role of Gravity Data in Determination of the Gravitational Potential of the Earth. Department of Geodetic Science, Report No. 134, The Ohio State University, Columbus, 1970.

Rapp, R.H., Improved Models for Anomaly Computations from Potential Coefficients. Department of Geodetic Science, Report No. 181, The Ohio State University, Columbus, 1972.

Rapp, R.H., Gravity Anomaly Recovery from Satellite Altimetry Data Using Least Squares Collocation Techniques. Department of Geodetic Science, Report No. 220, The Ohio State University, Columbus, 1974.

Reilly, J.P. and E.H. Herbrechtsmeier, "A Systematic Approach to Modeling the Geopotential with Point Mass Anomalies". Paper published in Journal of Geophysical Research, Vol. 83, No. B2, American Geophysical Union, February 1978.